HumRRO

Final Report
HumRRO FR-WATSD-99-23
May 1999

The Transition from KIRIS to CATS: Instruction, Communication, and Perceptions at 20 Kentucky Schools

Arthur A. Thacker
Lisa E. Koger
R. Gene Hoffman
Milton E. Koger

Prepared for:
The Kentucky Department of Education
500 Mero Street
Frankfort, KY 40601

Contract Number BP 010492

Human Resources Research Organization
66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314-1591

Acknowledgments

The authors of this report would like to thank the teachers, principals, and district representatives who took time out of their busy schedules to talk with us. Without their cooperation this study would not have been possible. We would also like to recognize the other school personnel and students for graciously welcoming us into their schools and classrooms. The office staff at each of the participating schools handled the lion's share of the coordination for interviews and observations and are to be commended for their excellent work. All of these persons helped us in our attempts to represent Kentucky's schools honestly and accurately within this report. Of course, any remaining shortcomings are the responsibility of the authors. The opinions expressed in this report are those of the authors and do not represent the opinion or policy of the Kentucky Department of Education.

THE TRANSITION FROM KIRIS TO CATS: INSTRUCTION, COMMUNICATION, AND PERCEPTIONS AT 20 KENTUCKY SCHOOLS

Table of Contents

List of Tables

THE TRANSITION FROM KIRIS TO CATS: INSTRUCTION, COMMUNICATION, AND PERCEPTIONS AT 20 KENTUCKY SCHOOLS

Abstract

During the spring 1998 legislative session, the Kentucky General Assembly passed HB 53, an amendment to the original Kentucky Education Reform Act of 1990. The old state assessment and accountability system, known as the Kentucky Instructional Results Information System (KIRIS) was replaced with a new system called the Commonwealth Accountability Testing System (CATS). KIRIS used rewards and sanctions to encourage improvements in instructional practices. This study was primarily interested in whether the change in testing systems from KIRIS to CATS would have an impact on these classroom instructional practices. Researchers also investigated how teachers learned about CATS and whether they thought it would be an improvement over KIRIS.

Researchers from the Human Resources Research Organization (HumRRO) visited 20 schools, 1 elementary school and 1 middle school from 10 school districts in Kentucky. The study was conducted before the first round of the Kentucky Core Content Tests (subject-specific tests that are one part of the overall CATS system) were administered in 1999.

Researchers collected data from classroom observations and from interviews with teachers, principals, and district representatives. They found that the transition to CATS has had little influence on the instructional practices currently in use in Kentucky classrooms. Most teachers reported that they made significant changes to their teaching practices when KIRIS began, and that they are continuing to use those practices this year as they prepare to give the Kentucky Core Content Tests for the first time. These practices, which emphasize application to real life, analysis, and evaluation, are viewed positively by most teachers in the study.

The research also showed that teachers rarely reported that they felt well informed about CATS. Very few teachers reported having any training concerning CATS, apart from a few who said they had been required to attend professional development sessions on "test administration ethics" or a symposium explaining the computation of the new school accountability index. The teachers reported that the ethics training was identical to training given under KIRIS.

Finally, most teachers reported that they were unsure if CATS represented an improvement over KIRIS since they had not seen the new test document yet. Many preferred to reserve judgment until after they received the first round of scores under the new system, and a sizeable segment expressed an element of cautious optimism about the CATS system because their faith in KIRIS was so low. They also said that they had little confidence that CATS would be fairer than KIRIS in the way it administered rewards and/or assistance to schools.

THE TRANSITION FROM KIRIS TO CATS: INSTRUCTION, COMMUNICATION, AND PERCEPTIONS AT 20 KENTUCKY SCHOOLS

Summary A
Instructional Practices During the KIRIS-CATS Transition

According to a report by the Human Resources Research Organization (HumRRO), the change in accountability systems from the Kentucky Instructional Results Information System (KIRIS) to the Commonwealth Accountability Testing System (CATS) in the 1998-99 academic year did not cause a great deal of change in instructional practice. HumRRO visited and collected data from 20 schools in 10 districts during the first year of a four-year study of the consequential validity of CATS (Thacker, Koger, Hoffman, & Koger, 1999) before the first administration of the Kentucky Core Content Test. Their findings show that the implementation of CATS has caused a ripple compared to the relative wave of influence the initial implementation of KIRIS caused. Many teachers are convinced that the new accountability system will not be different enough from KIRIS to justify significant changes in their practices. Numerous teachers quipped, "rumor has it that CATS is just KIRIS with a new name." Others contend that there is no way to judge what changes will be justified prior to seeing the test and the first round of results. They are content to "wait and see" at this stage.

The reform movement in Kentucky has generated considerable momentum during the past nine years. Reform-oriented instruction is common. Rote memorization from textbooks has been supplanted by attempts to access higher-order thinking skills. And, with very few exceptions, these changes are viewed positively by teachers. One teacher explained, "Ten years ago there was no guidance and no one to help. Now there are resource persons and performance standards that eliminate the easy way out, that is, always using puzzles and worksheets." Changes in instructional practice and the influence of those changes on KIRIS test scores are well documented (Hoffman, Harris, Koger, & Thacker, 1998; Harris, Hoffman, Koger, & Thacker, 1999). Kentucky's teachers have invested considerable time and effort learning to teach differently since KIRIS began. They are unwilling to abandon those practices without compelling evidence to suggest that they should do so.

CATS includes the Kentucky Core Content Tests, subject-specific tests of student achievement that will be used to incorporate test items in both multiple-choice and open-response formats in each school's accountability index. The KIRIS test also included both types of items, but the multiple-choice items were not used in the school accountability formula. This change in the accountability formula has garnered the most attention in terms of instructional practice, but the changes made with regard to including the multiple-choice questions are minor. Several teachers explained that they had always used multiple-choice questions to some extent and that the addition would not make a difference. Others claimed that they had begun to include more multiple choice items on their classroom assessments, but that their methods of instruction had not changed. One school's principal explained that the Scantron® machine the school had purchased just before the Kentucky Education Reform Act (KERA, the act that began school accountability in Kentucky) passed, was finally seeing some use.

CATS has had a similarly modest impact on teacher professional development. Teachers

and individual schools have a considerable amount of choice regarding their professional development (Thacker, Koger, & Koger, 1998), and they are typically not choosing to attend training specific to CATS. The few teachers who did report that they had received any training about the new accountability system had attended either a workshop about test administration ethics or a symposium explaining the proportion for which each part of the CATS accountability system will count. When asked to elaborate on the ethics training, teachers explained, "It's the same training we had for KIRIS." Most were more concerned with the allocation of points used to compute their schools' accountability index. The training they had received often left them with more questions than answers, especially regarding the norm-referenced portion of CATS.

School visits also pointed out several possible factors that may impact further evaluation research concerning CATS. Among those factors are the influence of teacher content knowledge on instructional practices, availability and implementation of reform-friendly teaching materials, self-contained versus departmentalized instruction in the elementary schools, schools' attempts to maximize student content learning in assessment grades, issues related to testing special needs students, and other issues. As CATS becomes more and more a part of the everyday language of schools and as schools strive to earn rewards under the new system, these factors have considerable potential to influence instruction.

Bibliography

Harris, C. D., Hoffman, R. G., Koger, L. E., & Thacker, A. A. (1998). *The relationship between school gains in 7th and 8th grade KIRIS scores and instructional practices in mathematics, science, and social studies for 1997* (HumRRO Report FR-WATSD-98-05). Radcliff, KY: Human Resources Research Organization.

Hoffman, R. G., Harris, C. D., Koger, L. E., & Thacker, A. A. (1997). *The relationship between school gains in 8th grade KIRIS scores and instructional practices in mathematics, science, and social studies* (LRS97-3). Frankfort, KY: Bureau of Learning Results Services, Kentucky Department of Education.

Thacker, A. A., Koger, L. E., & Koger, M. E. (1998). *Professional development under the Kentucky Education Reform Act: A study of practices in 30 middle schools* (HumRRO Report FR-WATSD-98-38). Alexandria, VA: Human Resources Research Organization.

Thacker, A. A., Koger, L. E., Hoffman, R. G., & Koger, M. E. (1999). *The transition from KIRIS to CATS: Instruction, communication, and perceptions at 20 Kentucky schools* (HumRRO Report DFR-WATSD-99-23). Radcliff, KY: Human Resources Research Organization.

Summary B
The Effectiveness of School Communication During the KIRIS-CATS Transition

What had teachers heard about the Commonwealth Accountability Testing System (CATS) prior to administering the Kentucky Core Content Test, a component of CATS? How did they learn about CATS and the Kentucky Core Content Test? These questions helped frame the first phase of a four-year study by the Human Resources Research Organization (HumRRO) concerning the consequential validity of CATS (Thacker, Koger, Hoffman, & Koger, 1999). HumRRO was interested in the possible influences the change from the Kentucky Instructional Results Information System (KIRIS) to CATS might have on instructional practices. In order to evaluate those influences it was important to establish how much the teachers knew about the changes. It was also important to recognize the sources of the information teachers did possess, both to evaluate the effectiveness of established communication channels and to identify the origins of possible misconceptions about the new accountability system.

In the course of visiting 20 schools in 10 districts around Kentucky, HumRRO found that teachers rarely reported that they felt well informed about CATS. When researchers asked teachers what they had heard about CATS, the most common response was, "Not very much." Very few teachers reported having any training concerning CATS. Those who did have professional development meetings about the new accountability system had only attended "test administration ethics training" or a symposium explaining what proportion each component of CATS would count in the computation of the school's accountability index. Those teachers who had attended the ethics training reported that it was identical to the training they had for KIRIS.

Teachers reported that they received their information about CATS from a variety of sources. Most commonly, they learned about the new system from materials placed in their school mailbox, from informal conversations with other teachers, and from the newspaper. Communication between teachers has dramatically improved during the past 10 years. The Kentucky Education Reform Act (KERA) and the associated KIRIS accountability system, stimulated a great deal of teacher interaction, primarily due to curriculum alignment efforts. Those lines of communication, as well as curriculum alignment efforts, are still very much in place.

Teachers often reported that their information came to them in the form of memoranda in their school mailbox. They were often unsure of the origins of these documents, although many assumed that they came from the Kentucky Department of Education (KDE). When we spoke with the principals of the schools, however, we learned that the district office was much more likely to have provided the information. In a way the teachers were not incorrect, because KDE does rely on the District Assessment Coordinators (DACs) as a primary channel of communication to the schools.

Relying on the DACs for getting information to the schools, we discovered, has the

potential for two very different problems. The first is that the DAC's time is often very limited in small rural school districts. Schools in these districts typically send representatives to the DAC to collect information for the rest of the school's personnel. Each reinterpretation adds to the possibility that the information will become diluted or altered in some significant way. When schools have specific questions or issues to be addressed in these districts they are often forced to wait for the availability of the DAC. The other end of the spectrum exists in large urban districts. These districts produce so much material for the schools that teachers often do not have the opportunity to sort and interpret the information they receive. Whatever the communication problem, be it a lack of sufficient material or an overabundance of it, the flow of communication from KDE to teachers remains effectively stifled.

KDE maintains a web site on the Internet, complete with e-mail addresses to which teachers and schools can address their questions; however it was rarely if ever mentioned by teachers as a source of information about CATS. The KDE newsletter, the *Kentucky Teacher*, was only very rarely mentioned. Teachers seem to depend on their school to provide them with the information they need. Directly mailing schools may be a more effective choice for distributing important information to teachers than the various other more complex methods currently in place.

Bibliography

Thacker, A. A., Koger, L. E., Hoffman, R. G., & Koger, M. E. (1999). *The transition from KIRIS to CATS: Instruction, communication, and perceptions at 20 Kentucky schools* (HumRRO Report DFR-WATSD-99-23). Radcliff, KY: Human Resources Research Organization.

Summary C
Teacher Perceptions of the Value of the New Testing Program

How confident are teachers that the Commonwealth Accountability Testing System (CATS) represents an improvement over the previous Kentucky Instructional Results Information System (KIRIS)? Researchers from the Human Resources Research Organization (HumRRO) interviewed teachers from 20 schools in 10 districts regarding their confidence in CATS (Thacker, Koger, Hoffman, & Koger, 1999). Although the majority of those teachers could recount the substantive changes to the accountability system, very few stated that they were convinced that the system would be better. A large number of the teachers said that the new accountability system represented a change in name only.

The most common response from teachers was that the new system was still largely unknown. They said that before they could make sound judgments about its worth, they would need to see the test and receive the first round of student scores. The next most common response from teachers can be classified as cautious optimism. The actual responses ranged from "It can't be worse than KIRIS," to "Just adding the multiple-choice questions would make the test less subjective and that would make it better in my opinion."

Another common response from teachers was to relate an opposition to the accountability system on a philosophical level. For these teachers, substituting one reward system with another makes very little difference. They claim that by emphasizing monetary rewards, the system-whether CATS or KIRIS-adds a negative connotation to teaching. "Any time that money is involved someone will find a way to cheat or to play the system," said one of the participating elementary teachers.

Teachers are also not convinced that the change in accountability systems is going to have an effect on the preparation of their students for the next grade level. The overwhelming majority of teachers said "not any" when asked about the influence of CATS on their students. A few teachers said that the addition of multiple-choice questions might give the students some needed practice with the format, which in turn, might help them on other multiple-choice format tests. None professed that their student would know or be able to do more as a result of the change in accountability systems.

Teachers were also asked if they thought that the new system would be fairer than the old one. The majority had very little confidence that the CATS system would help the state administer rewards or assistance more fairly than KIRIS. A few teachers applauded the change in the mechanics of administering rewards from giving rewards as bonuses to administering them as school funds, but they were not convinced that the new system would be better at determining which schools received the rewards in the first place.

House Bill 53 suggests placing students' Kentucky Core Content Test scores (a component

of CATS) on their transcripts. Teachers supported this effort toward student-level accountability, but also said they believe that it would make very little difference at the elementary and middle school levels. “It might mean more to high school students, especially if the colleges and universities use the scores to determine admittance or eligibility for scholarships.” They were also quick to point out that the students who would be most affected by the addition of scores to their transcripts are the students who are already highly motivated to do well on the test. “The students that don’t care about things like grades and transcripts are the ones that need the most attention.”

When teachers and principals were asked about their use of the KIRIS score reports from last year, most said that they were used much the same as in previous years. The scores are used as a diagnostic tool for the preparation of specific programs and policies that address each school’s improvement goals. Even though KIRIS might be considered a “lame duck” this year, the scores from last year are being examined closely. Schools used those scores to plan professional development and design their Consolidated Plans (KDE, 1997). When principals discussed the scores it became very obvious that the practitioners of public education have not internalized the break between KIRIS and CATS yet. They expect to be able to compare last year’s KIRIS scores to this year’s Kentucky Core Content Test scores. Many have programs in place that may be bolstered or eliminated because of perceived changes in student scores. The *Kentucky Teacher* (Fishback, 1999) also suggests that attributing growth will be possible using both KIRIS and CATS scores. The words “interim period” were rarely heard during interviews.

Bibliography

Thacker, A. A., Koger, L. E., Hoffman, R. G., & Koger, M. E. (1999). *The transition from KIRIS to CATS: Instruction, communication, and perceptions at 20 Kentucky schools* (HumRRO Report DFR-WATSD-99-23). Radcliff, KY: Human Resources Research Organization.

THE TRANSITION FROM KIRIS TO CATS: INSTRUCTION, COMMUNICATION, AND PERCEPTIONS AT 20 KENTUCKY SCHOOLS

Introduction

Background

With the passage of the Kentucky Education Reform Act (KERA) in 1990 (Kentucky Revised Statutes 158.645), Kentucky educators were forced to look upon teaching and learning in ways that were different from what had routinely taken place in the state's classrooms. Under this act, students were to be taught more than how to recall basic facts and dates; instead, the emphasis would be on being able to use knowledge in situations that required students to think critically and creatively and to do so at higher levels than previously required (Kentucky Department of Education, 1995).

The Kentucky Department of Education (KDE) was charged with developing a system that would achieve these ambitious goals by stimulating changes in classroom instruction, management, and funding allocations. The resulting plan called for the creation of a statewide assessment, known as the Kentucky Instructional Results Information System (KIRIS), which would address the issue of stimulating changes in instruction. This test, originally given to 4^{th}, 8^{th}, and 11^{th} grade students, required that students be able to explain their answers and to show what they were able to do with their knowledge. Instead of focusing only on multiple-choice or fill-in-the-blank questions as tests had in the past, KIRIS assessed students in several novel ways. For example, students answered open-response questions that required a written response of several sentences; they created writing and mathematics portfolios; they took part in performance events, in which small groups of students would work together to solve a problem and then write their solutions individually; and they responded to on-demand writing prompts.

To ensure that schools were indeed teaching toward these higher level skills, KERA also built into the system a means by which schools would be held accountable for their students' performance on the KIRIS tests. Simply stated, schools were given 20 years to achieve a target goal; those that exceeded expected performance on interim goals would be rewarded monetarily; those that failed to progress were sanctioned. These rewards and sanctions occurred at the end of two-year testing cycles. The "in crisis" designation meant that parents would be permitted to withdraw their students from the school and enroll them in schools that were making better progress toward the goal. The most severe sanction included the dismissal of school staff members and having the administration of the school taken over by state appointed representatives. However, confidence in the accountability system's validity soon eroded and this most serious step was never implemented. Subsequent research found that, for teachers, the threat of sanctions was a bigger motivator than the possibility of rewards (Abelmann & Kenyon, 1996; Kelley & Protsik, 1996). Sanctioned schools often received the assistance of Distinguished Educators, who were expert teachers specially trained by the state to work with these problem schools. In addition, schools that failed to improve were required to submit school transformation plans, which listed specific steps the school planned to take in order to ensure improved student learning.

School accountability for student learning was just one piece of the KERA "pie." Other important "pieces" included the issue of school improvement in areas such as funding, management, and teacher professional development. Management of the school, for example, became the responsibility of an elected committee of teacher, parent, and administrator representatives. This committee, known as the Site-Based Decision Making (SBDM) council, was given responsibilities ranging from hiring staff to selecting teacher professional development to submitting grant and funding proposals. Large disparities in funding between wealthier and poorer districts, the issue upon which the original KERA legislation had been based, were dramatically reduced.

The Evolution of KIRIS

Such a broad instrument for change quickly became a target of criticism from teachers, politicians, and concerned citizens with claims that the KIRIS test was neither reliable nor valid, despite a favorable independent review of the test published by The Evaluation Center (1995). The release in 1995 of the Hambleton et al. report stating that KIRIS was "seriously flawed" added additional fuel to the fire. Later that year, state education officials accepted 10 of 12 changes to the testing system recommended by the General Assembly's Office of Education Accountability. These recommendations, to take effect in the spring 1997 round of KIRIS tests, included adding multiple-choice questions to the test, dividing the test among more grade levels, eliminating mathematics portfolios and performance events from the assessment, and adding a norm-referenced test for 3rd, 6th and 9th graders. However, because multiple-choice questions had to be developed and field tested, these questions would not count towards accountability scores until spring 1999 testing.

An additional criticism—that teachers did not know what parts of their content area would be considered "fair game" on the assessment—was addressed with the release in summer 1996 of the *Core Content for Assessment* document (KDE, 1996). This document specified what was eligible for inclusion on the KIRIS tests, both by grade levels and by topic. Prior to its release, teachers had only the six Learning Goals developed by the Kentucky General Assembly and the 57 Academic Expectations with which to develop their curriculum. These goals and expectations still provide the final guidance on what Kentucky students should know and be able to do when they graduate from high school. Both the Learning Goals and Academic Expectations can be found in *Transformations: Kentucky's Curriculum*, released by the Kentucky Department of Education in 1995. The Program of Studies and the Program of Studies Implementation Guide (KDE, 1998b) are further refinements/expansions of *Kentucky's Core Content for Assessment*.

Other changes were made over the years, as well. The 1998 score reports, for example, distinguished among three levels of both novice and apprentice student-achievement categories. Previous score reports had categorized just one level of novice and one level of apprentice students. This made it difficult, said some teachers, to determine whether those lower achieving students were making any progress. It was especially troublesome for special education students to show progress under the old system, teachers said. The majority of Kentucky's students have fallen into either the Novice or Apprentice categories since the beginning of the accountability system. These additional scoring levels for novice and apprentice will be computed in the new accountability formula.

Finally, the concept of transformation planning changed to that of consolidated planning with the release in 1997 of the *Kentucky Consolidated Planning Process* (KDE, 1997). This document provided guidance for school staff in the funding and grant writing process. Under consolidated planning, the funding of school activities and programs would have to be done as a collaborative process, rather than the often piecemeal way in which transformation planning had previously been done. The state's Distinguished Educators, who were assigned to schools not showing progress, were trained to assist in the consolidated planning process, as well.

The Demise of KIRIS and the Development of CATS

Despite the changes to the original testing system, controversy and criticism continued to surface. In June 1997, for example, the testing contractor, Advanced Systems in Measurement & Evaluation, Inc., was fired after an error was discovered on the scoring of 1996 tests. The error later caused scores of nearly 400 schools to be readjusted after being scored too low the first time through the scoring process (Lawson, 1997). Another independent evaluation of the testing and accountability system revealed many areas of concern about the validity and reliability of KIRIS and the fairness of the accountability system (Koretz, Barron, Mitchell, & Stecher, 1996).

The Kentucky General Assembly re-examined the state's testing program during its spring 1998 session. After several bills in both the Senate and the House were considered, the General Assembly passed House Bill 53, which became law on April 14, 1998. No longer would the state's testing system be known as KIRIS; instead, it would be called the CATS (Commonwealth Accountability Testing System), and instead of the KIRIS test, students would now take the Kentucky Core Content Test. The new test would implement recommendations already made by the Office of Education Accountability as well as other changes:

- Multiple-choice questions, which had appeared on previous KIRIS tests but had not counted toward a school's score, would count on the 1999 Kentucky Core Content Test, along with open-response questions.
- The test would continue to be split among 4^{th}, 5^{th}, 7^{th}, 8^{th}, and 11^{th} graders, but 10^{th} grade students would also be tested for the first time on the practical living/vocational studies and reading test sections. The writing prompt also moved to the 12^{th} grade.
- The norm-referenced test given to 3^{rd}, 6^{th}, and 9^{th} grade students would continue to be given, and it would be included in the accountability formula. This was an important consideration for those who had criticized KIRIS for not having an established method in place to compare Kentucky's students to those in other states.
- Money previously given directly to teachers at schools in reward status would now go to the school. Under the previous system, teachers at schools in rewards reported being stressed by having to decide who was eligible to receive rewards and how much they were to receive (Abelmann & Kenyon, 1996; Kelley & Protsik, 1996).
- The number of required writing portfolio pieces would be reduced, from six to four for elementary students and to five for older students. The time spent on portfolios, both in writing and grading, was a major complaint of teachers under the previous system.
- The test itself was streamlined, with responses on open-response questions limited to one page. This, it was hoped, would reduce the amount of time necessary to complete the test,

reducing stress on students and increasing the time allotted for instruction.

- Results were to be returned by September 15, in order for schools to use those results in planning curriculum and instruction for the next group of students who will take the test. Previously, the lag in time from testing to reporting results made it difficult for teachers to adjust their curriculum (Thacker, Koger, and Koger, 1998).

In September 1998, CTB/McGraw Hill won the contract for developing the CATS test, with these requirements (for elaboration or clarification, see the Kentucky Department of Education's Request for Proposal, 1998a).

The Consequences of Implementing CATS

There are both intended and unintended consequences of implementing any test. Although many teachers contend that "teaching to the test" is a bad thing, we nonetheless recount the old adage that "what gets tested, gets taught." In Kentucky's case, this can be changed slightly to "how something gets tested becomes how something is taught." In other words, it would be strange that a test that emphasizes writing would not have any impact in the routine teaching and learning that is taking place in the classroom. Therefore, it would be expected that Kentucky's KIRIS tests, which required written explanations on how one solved a problem, would influence classroom instruction as well. In fact, two previous studies (Hoffman, Harris, Koger, & Thacker, 1997; Harris, Hoffman, Koger, & Thacker, 1998) showed not only that KERA stimulated reform practices, but that those practices were positively correlated to KIRIS gains.

Other intended consequences of KIRIS, and now CATS, influence the professional development taking place in schools and the emphasis being placed on curriculum alignment. Thacker, Koger, and Koger (1998) examined the professional development issue in a study of 30 Kentucky middle schools. They found that teachers generally said that professional development since KERA had changed for the better. Specifically, teachers said they had more choice in selecting professional development that would directly benefit their students; they also said that professional development committees selected professional development that targeted their schools' weaknesses, which were uncovered by examining test results. Long-term teachers recounted examples of professional development from their early days in which all teachers were required to attend district-wide inspirational rallies that they described as having little or no benefit. On the other hand, some teachers complained about the narrow focus that their professional development had assumed, especially if they were in a declining school. In those instances, teachers had little choice in their professional development, which was prescribed by their Distinguished Educator with the intent of raising their scores.

In the same study of 30 Kentucky middle schools, Thacker, Koger, and Koger (1998) also examined curriculum alignment as an intended consequence of the KIRIS test. Curriculum alignment ensures that gaps (the overlooking of required topics in life science, for example) and duplications (studying World War I in both seventh- and eighth-grade social studies was one particular example) in curriculum are eliminated, resulting in a smooth continuum of instruction throughout a student's educational career. The authors found that teachers from high-gaining schools reported more alignment efforts at the district or county level, rather than just at the school level, than did those teachers from low-gaining schools. Teachers from nearly all schools visited reported that at least some curriculum alignment had been accomplished, although some were

farther along than others in the process.

An example of an unintended consequence of the testing system was the apparent over-interpretation or misuse of released items from the KIRIS test (Thacker, Hoffman, & Koger, 1998). During workshops held with either middle school science or social studies teachers, it was discovered that many teachers were using the released items to design their curriculums. Despite the fact that items were only released once they were taken out of the active item pool, teachers spent a great deal of time preparing students to answer those specific items. Several teachers described elaborate "roller-coaster units" designed around a particularly problematic science item. For many teachers, the released items represented the most obvious link to the test itself. By teaching the released items those teachers apparently hoped that they would "get lucky" and hit the exact content included in the items on the upcoming test. They regarded the match between their teaching and the tested content to be largely a matter of chance. The concept of teaching students to generalize from the concepts included in the *Core Content* to the specific questions on the test was either considered infeasible or not considered at all.

A common complaint of teachers about the influence of KIRIS was that they were forced to neglect the "basics" of their subject—mathematics computation or writing mechanics, for example—in order to teach the "process" skills that KIRIS demanded. These process skills generally are defined as what one "does" with the basic facts—thinking at higher levels (analysis, synthesis, evaluation), applying facts to a real life situation, problem solving, or writing. In particular, teachers and others have been quick to declare that students' spelling ability, along with other aspects of writing mechanics, have declined since the implementation of KERA. They claimed that they were unable to focus on these skills because they were being forced to spend inordinate amounts of time teaching their students how to write the responses now required under KERA. To test this claim, Hoffman, Koger, and Awbrey (1997) conducted a study in which 1993 KIRIS writing prompts from fourth-grade students were compared to identical prompts from 1996 fourth graders. The prompts were scored, with the assistance of elementary school teachers, on the basis of the number of spelling and grammatical errors (for example, subject-verb agreement or run-on sentences). The researchers found that, instead of declining, girls' writing mechanics over the four-year period had remained unchanged, while boys' writing mechanics had actually improved.

With the transition to the new CATS testing system, one can expect other consequences to come about. Some can be anticipated; some, on the other hand, will fall into the "unintended/unexpected" category. It is reasonable, for example, to expect that teaching practices may change as the result of the switch from an all open-response/writing format testing system to one that combines writing and multiple-choice questions and which includes a multiple-choice norm-referenced test.

Other changes will be apparent in the accountability system, which determines whether rewards or assistance is given to schools. The new system includes several important changes:

- The system will permit additional levels of novice and apprentice students. Previously, novice students earned a "0" and apprentice students received a "40," but the new system has added scores of 13 and 26 in the novice category and 60 and 80 in the apprentice category.

- A simplified method of determining the success or decline of schools will be set in place. A "straight line" will be drawn from a school's baseline score to its Year 2014 goal of 100 (proficient performance), thus indicating the amount of growth a school must achieve over the years. Schools meeting or exceeding their goal would receive rewards. An important addition to the "straight line" method is the addition of a "safety zone" that in effect protects a school from small declines that may be attributed to the margin of error in the school's scores. Schools dropping below the safety zone would come under state scrutiny and intervention (Harp, 1999). This new system is vastly different from the previous one, in which goals were recalculated every two years. Schools that failed to meet their goal even slightly were sanctioned. As mentioned previously, several studies have shown that fear of sanctions was a bigger motivator than the prospect of receiving rewards (Abelmann & Kenyon, 1996; Kelley & Protsik, 1996). The reduction in the likelihood of receiving sanctions may have some important consequences in the future. In addition, the fact that schools rather than teachers will receive rewards may reduce stress among staff, who often found themselves embroiled in battles over who should receive portions of the reward money (Abelmann & Kenyon, 1996; Kelley & Protsik, 1996; Holland, 1998). It might also affect teachers' motivation to change.
- The weights given to different portions of the accountability system have been changed. Previously, for example, all core content scores counted equally for elementary schools, but beginning in 2000 certain subjects will account for different portions of the accountability total. Mathematics and reading each will account for 19% of the accountability "pie" with science, writing, and social studies each receiving 14.5% and arts/humanities and practical living/vocational studies each receiving 4.75% of the total. Also, the norm-referenced tests that have been given to third-, sixth-, and ninth-grade students will be included in the accountability formula for the first time. They will account for 5% of a school's total (Fishback, 1998/9). These recomputations have the possibility of influencing instruction. Elementary teachers may devote more instruction time to mathematics and reading, for example, since they will account for a greater percentage of the accountability total. And with the addition of the norm-referenced test to the accountability formula, it is likely that many teachers will devote more instruction to the basic skills that are supposedly measured by this type of test.

A good deal of information is being released about the transition in the *Courier-Journal*, the *Herald Leader*, and other local newspapers (news items). Also much effort is being spent through KDE vehicles such as the *Kentucky Teacher* as well as the recently revamped KDE web site. DACs are acting as liaisons between KDE and the schools within their districts.

Primary Areas of Study

CATS is designed, as was its predecessor, to stimulate change in Kentucky classrooms. Its purpose is to improve student achievement through improved instruction and improved student motivation. There are several methods built into the CATS system that are designed to effect this desired change:

- Establishing clear student goals
- Use of "active learning"
- Connection of knowledge through multiple assessment formats
- Rewarding successful schools
- Providing meaningful student reports and possibly scores on students' transcripts

Schools are further guided to improve by the accountability system and its accompanying documentation, such as annotated released items, the Interpretive Guide, and score reports for district, school, and student levels.

This report represents the first phase of a four-year, four-phase program of research examining the validity and consequential impact of CATS. Two methods of data collection, school visits and teacher surveys will be utilized in these studies. HumRRO will coordinate with KDE on the preparation of the teacher survey and share analysis and reporting of survey results with KDE. The school visits portion of the research will address the validity issue of consequential impact by examining the impact of CATS, in general, and the Kentucky Core Content Test, in particular, on instruction. Analysis of the teacher questionnaire will address relationships between teachers' descriptions of their instruction and CATS results. HumRRO's research will address science and social studies in elementary and middle schools.

This data collection and analysis allows for considerable continuity across the entire four years of study. With the number of schools we are slated to visit, qualitative methodology will likely prove more fruitful than a quantitative approach. It is easy to imagine the visits studies as "school stories" research. Each year, through teacher interviews, classroom observations, and artifact analysis, HumRRO will tell school stories in ever-increasing detail. Those stories will then be examined in reference to quantitative measures of school practice in the form of the student and teacher questionnaires and the scores themselves. The degree to which these stories are similar or different from each other, in reference to the other measures, should serve to inform educators about the effectiveness of various educational policies and programs at the school level.

This plan represents a compromise between quantitative and qualitative research. The teacher survey data should be sufficiently reliable to satisfy generalizability concerns, while the visits themselves give the study depth and frame the findings of the survey research in terms of what occurs in the everyday practice of teaching and schooling (Thacker & Hoffman, 1998).

The purpose of this first phase of the research was to examine the impact CATS is having thus far on Kentucky's teachers and administrators—and by extension on Kentucky's students. If CATS is designed as an instrument of change, is that change being implemented in the state's schools? We focused upon three major areas, instructional practice, communication, and

confidence/perceptions. Research questions were developed from these issues. These questions provide the structure of the study:

Table 1 Research Questions by Instructional Practice, Communication, and Perceptions

Instructional Practice	
1	What changes in classroom instruction have taken place as a result of the change in testing systems from KIRIS to CATS?
2	What types of instructional practices are being used in classroom instruction?
3	Has the change in accountability systems affected the level or quality of work that is expected from students?
4	What influence has the addition of multiple-choice questions had on instructional practice?
5	What influence has CATS had on teacher professional development?
Communication	
1	What have teachers heard about CATS?
2	How have teachers learned about CATS?
Perceptions	
1	Does CATS represent an improvement over KIRIS?
2	Will CATS affect student preparation for middle and high school?
3	Is the reward system fairer under CATS than KIRIS?
4	What effect will putting CATS scores on student transcripts have?
5	How useful were the 1998 KIRIS score reports?
6	How could score reports be improved?

Methodology

Case Study Description

This research is best defined as a case study. It is not typical of case study research due to the implicit need to generalize from the schools that were chosen to the larger system. Generalization like this is not typical of case study research but applies for this particular study if we examine the caveats and limits of the study. First, the case that was studied was the Commonwealth Accountability Testing System (CATS), and not any of the schools that we visited. The schools represent the primary level of actors that were observed to learn about the case selected. The secondary levels of actors were the classes within the schools and the tertiary actors were the many teachers, administrators, and district personnel we interviewed. In order to learn about the system, we studied the schools, the classes, and the people who comprise them. This study can be compared to a more typical education case study where researchers study classes,

principals, teachers, and students in order to learn about a school. Our study simply started one level higher and evaluated the system itself.

Second, the study was evaluative of the case, but not in traditional case-study terms. Instead this study examined an aspect of consequential validity, as defined by Messick (1989). In short, this study sought to discover the impact CATS had on the everyday practices of teaching and learning in public schools. Particularly, this study sought to determine the impact of changing the accountability system from KIRIS to CATS on classroom practices.

This idea leads to the necessity for generalizations to be made from the results of this study. The common problem with generalizing from case study research is that often one case does not act like another. In this instance another case would be another state's accountability system, and that is not the level of generalization with which we are concerned. In case study research it is common to find that actors within a case will react to external factors in similar ways (Stake, 1995). For example, students will tend to exhibit certain behavior patterns within a specific school environment. We might therefore expect schools to exhibit certain patterns within the CATS system. Generalization from these actors within CATS may tell us a great deal about the expectations and reactions of the larger system.

Selection of Sites

Several criteria were used to select the 10 districts for this study:

Geography

Geography was the most important criterion for selecting schools. To represent districts from all over the state, schools in 8 of the 9 regional service areas were visited (Only one district refused our request to participate, and it was in the one region we failed to visit. Schools from this region will be included in Phase Two of this study, which will begin next year.). Two regional service centers each had two districts participating in the study. These districts were chosen to ensure that both rural and urban districts were adequately represented.

Size

A size criterion was established to ensure that extremely small schools were not included in this study. Generally, we visited schools that had a student population of at least 100 students per grade level. This was to help reduce the impact that a small sample size might have on test scores, size of teaching staff, available resources, etc.

Separate sites

Nine of the 10 districts had "freestanding schools," for lack of a better term. This means that elementary students were housed in their own building (K-5 or K-6); middle schools, with one exception, were also housed separately in 6-8 or 7-8 configurations. Our exception was a middle school that shared a building with the high school; however, the middle and high schools functioned separately from each other, each with its own administration, staff, and section of the building.

Recent KIRIS scores

In this study, past performance on KIRIS was not a major selection criterion, but it was considered. Schools within regions/districts were chosen to exhibit a wide range of KIRIS scores. Some schools had received rewards, while others had been in decline. Still others had bounced from one extreme to the other.

A list of possible participant schools was developed from these criteria. This list was sent to KDE, who sent letters (Appendix A) to the candidate schools and their districts describing our study and inviting their participation. Follow-up phone calls to districts and schools were made to answer questions and schedule visits. As mentioned previously, only one middle school refused our request to participate in the study. Since it was the only middle school in that district, a replacement district was necessarily selected.

After schools agreed to take part in the study, they received information packets to help them prepare for the site visit. Elementary and middle schools received nearly identical packets, with only minor editing differences to make them suitable for the particular situation (both elementary and middle school documents are included in the Appendices section). The packets contained:

- A letter to the principal confirming the visit and enlisting his/her help in preparing for the research team (Appendix B)
- A form to be used in scheduling interviews and observations (Appendix C)
- A form to be returned to the research team prior to the site visit listing motel information and the school's location (Appendix D)
- Information letters to be given to teachers describing their role in the research effort and enlisting their cooperation (Appendix E)

Coding and Interpreting Data

A group of four researchers conducted the site visits. A two-person research team visited each district for about three days. Typically, this visit occurred on a Monday, Tuesday, and Wednesday and the researchers used the remainder of the week to write notes, interpret and organize data, and prepare for further school visits. Schedules were occasionally altered at the behest of the schools or to eliminate conflicts with school activities and events.

The teams were reconstituted often to minimize observer effects. In addition, all researchers conducted a mock data collection at a convenient public school that was not asked to participate in the study. The teacher interview schedule (Appendix F) was fine-tuned using the field tests. Researchers used a consensus-building exercise in order to calibrate classroom observations. From these discussions, it was decided that the interaction between the teacher and students was the most robust descriptor of the classroom. All researchers began their observation notes by describing this dynamic. Details were added to the observation notes as they became important and/or obvious. This methodology produced similar observation notes in our calibration sample. During data collection researchers frequently observed the same classes as a further check

of consistency. More global school-level descriptors were also discussed to ensure that the portrayals of the participant schools were genuine and unbiased.

In addition to conducting teacher interviews and observations, researchers interviewed each principal (Appendix G) and a representative from each district (Appendix H), generally a District Assessment Coordinator or Instructional or Curriculum Supervisor.

Results

Demographics of Schools

Table 2 lists pseudonyms for the participating schools. It also includes a short description of the school that indicates whether the school is rural, suburban, or urban, where the school is located in Kentucky, and a very short identifying comment. This table makes no attempt to fully describe the schools. It is provided as a reference due to the relatively large number of schools mentioned later in the report. A more complete description of the schools follows. The table demonstrates the variety of schools contained within the sample group.

Table 2 Participant School Descriptors

District	*School*	*Descriptors*
Elm County	Elementary	Large, new, rural, Eastern
	Middle	Large, competitive, rural, Eastern
Oak Independent	Elementary	Suburban, writing emphasis, Central
	Middle	City school with some urban aspects, 7th and 8th grade only, Central
Pine	Elementary	Urban, diversity requirements (busing), Central
	Middle	Urban, poor academic reputation, high percentage of special education students, Central.
Spruce	Elementary	Rural, returning to more traditional teaching after reform emphasis, Western
	Middle	Rural, technology emphasis, Western
Cedar	Elementary	Suburban, older school, experienced staff, Northern
	Middle	Suburban, extreme teaming strategy and schedule, Northern
Hickory	Elementary	Rural, small town school amid several other "county" schools, Central.
	Middle	Rural, small town school, Central.
Poplar	Elementary	Rural, diverse population, switching 3rd and 4th grade teachers, Western.
	Middle	Rural, dropping student population, Western.
Cottonwood	Elementary	Rural mountains, mining community, Eastern.
	Middle	Rural mountains, double classes in assessment grade subjects, Eastern.
Locust	Elementary	Urban, innovative, older building with large classrooms, Central.
	Middle	Urban, contesting KIRIS scores, Central.
Walnut	Elementary	Rural, K-6, returning to basics, Southern
	Middle	Rural, 7-8, Southern.

Elm County School District

Elm County is a rural school district located in the hills of Northeastern Kentucky. The students were generally pleasant and polite. The teachers were friendly, although a little nervous about being observed and interviewed. Most of the employment opportunities in the area were either in the service industries or nearby factories. The area has recently seen a reduction in the number of available factory jobs, creating some changes in the economic demographics of the school. The effects of these economic changes have only recently become manifest. Chapter 1 funds are now available, and one of the science teachers explained that she no longer required students to complete a science fair project because many of the students could not afford to construct one.

Elm County Elementary School

Three elementary schools were consolidated into the newly built Elm County Elementary School. The facility is large, and only about five years old, but staff are already complaining that it is too small for the more than 750 students who attend. The classrooms are a bit small for the number of students and the amount of materials present. A couple of teachers suggested that the space occupied by the elaborately decorated courtyard in the center of the school would have been better used had it housed classrooms.

Despite becoming a Chapter 1 school recently, Elm County Elementary has been a reward school for the past two accountability cycles. The Chapter 1 funding has allowed them to hire additional teachers who assist with preparing students using reform pedagogy. One of these teachers serves as a writing specialist and works with regular classroom teachers by conducting guest lessons or teaming opportunities. Most of the classrooms had posters dealing with writing mechanics. The school has an extended writing class each day that lasts approximately one and one-half hours. Another of the "Chapter teachers" serves in more of an administrative role.

Teachers at Elm County Elementary seem worried about the recent loss of experienced teachers and staff. The school's principal was also new. He was the assistant principal previously. Ten teachers took early retirement last year. Evidently the district offered some incentives last year that made early retirement attractive. Remaining teachers believed that those who left were burnt out or frustrated with the accountability system.

Elm County Middle School

Elm County Middle School is located less than a mile from the Elementary school. It is also a relatively new facility. It serves about 800 students and is designed for a team-teaching approach. The grades are divided by floors and divided again into teams who inhabit color-coded halls. Each team has a common planning period during which students attend enrichment classes, such as music, art, physical education, etc. During the enrichment periods the entire hall (or all of one team) nearly empties of students. Researchers saw teams meeting with parents or making copies for future lessons during these planning periods and we were told about thematic unit planning, designing accommodations for special education students or other purposes.

The principal was very enthusiastic about block scheduling, which was implemented for

the first time this year at the middle school. Teachers were also very positive about the new schedule. No one voiced a dissenting opinion during the visit. Priority is given to assessment grade subjects in the blocking scheme, so seventh-grade science and eighth-grade social studies meet every day. Off-assessment grade subjects meet every other day. So eighth-grade science meets every Tuesday and Thursday plus half of all Fridays. This particular type of block will be altered next year due to concerns of some of the off-assessment grade teachers. Next year 75-minute blocks will replace the 90-minute blocks and core classes will be held in all grades daily.

One striking aspect gleaned from talking to the staff at Elm County Middle was the competition between it and the other county middle school. These two nearly identical schools compete in almost every aspect of public education, from sports scores to test scores. Elm district has two similar high schools as well, so the competition level continues even after the students leave the middle school. The second striking aspect of the Elm schools was the homogeneity of the student population. No ethnic diversity was evident at either the elementary or middle school.

Oak Independent District

Oak Independent School District is located in a mid-sized city. The area appears to have a broad range of employment opportunities, from service and factory jobs to professional activities; however, some job losses in the community have hurt local schools. The community is also quite racially segregated according to the principal of Oak Independent Junior High School. This segregation is reflected in the student enrollment of several of the district's elementary schools. The principal noted, "Some students have had little contact with students of another race before entering the junior high school as seventh graders."

Oak Independent Elementary School

Oak Elementary School is located in a comfortable middle-class neighborhood. It is an older school that has been updated and added to several times over the years. A neighborhood very near the school contains homes in the $300K to $400K range; generally, these homes are in the county district, but many parents pay tuition to have their children attend Oak Elementary. The principal told us that her student population of slightly more than 400 is more ethnically and financially diverse than the surrounding neighborhood would suggest. While some tuition students live in the wealthy neighborhood previously described, others are bused from city housing projects. The principal is worried that the opening of a county elementary school may have a negative impact on her school's demographics. Some of the county district parents who currently pay tuition to send their children to Oak Elementary may choose to send their students to the new school instead, she said. This possible change in demographics is causing her some concern over the impact this may have on their scores, and she expressed some concern about comparing two different groups of children to each other. Factory closings also have resulted in transfers, with the loss of additional middle-class students at this school.

This school has a strong base of parental involvement. The principal says that parent volunteers are "instrumental" in operating their computer lab and the Accelerated Reading program. The PTA has also raised thousands of dollars for various school projects.

Oak Elementary has a strong writing emphasis, which was obvious in several of the classrooms we observed. One teacher, for example, set high expectations by showing a

"proficient" writing sample to her students and then asking for suggestions on how to improve it. The writing sample was a test rubric example of what proficient writing looks like. The teacher was clearly not accepting proficient as the limit of her students' performance. Instead, these students were beginning with proficient writing and striving toward distinguished. Writing samples from students, school staff, and parents are also displayed on a "Wall of Writers" in the main hall near the office.

The interviewed teachers genuinely seemed to feel that they were on the right track in the way they were teaching. They had experienced success in the assessment system in previous cycles. Unlike at some other participant schools, no frustrated teachers were encountered at Oak Elementary.

Oak Independent Junior High School

Oak Junior High School is a seventh- and eighth-grade school of about 500 students, about one third of whom are minority students. As at Oak Elementary, a sizeable number of students—about 20%—live in the county school district but pay tuition to attend Oak Junior High. About 40% of the students at this school are described as being "seriously skill deficient" by the principal. He also describes many of his students as "homeless"—they have houses but not "homes" in the traditional sense of the word. Many, for example, are left to prepare their own dinners or otherwise care for themselves. Homework completion is a problem. Formerly, the school had been in decline; now, however, the school is in Rewards 2 status.

The facility consists of an older original building that has been added to over the years. The principal says the district is building a new middle school set to open in a couple of years. When that happens, the sixth grades from the district's elementary schools will be moved to the middle school, a move which he welcomes. He says that will eliminate some communication and curriculum alignment problems that the junior high currently experiences with some of its feeder elementary schools. He also says that his staff has invited sixth-grade teachers from the district's feeder elementary schools to attend meetings at the junior high, but only a couple of elementary schools have chosen to participate.

Oak Junior High tries to meet the needs of its diverse student population in several ways. The school offers three different honors classes per grade (science, language arts, and mathematics for seventh graders and social studies, language arts, and mathematics for eighth graders), and students can be placed in them according to their individual needs. A student may take honors science in the seventh grade, for example, without being automatically placed in honors mathematics. Students used to be more strictly tracked along ability levels, according to the principal, but school officials have reduced the amount of tracking somewhat.

The school also requires a series of enrichment classes for all students given before the start of the regular school day. The enrichment classes target assessed subjects by year; seventh graders take rotating week-long mini-courses in science and language arts and eighth graders study mathematics and social studies. These mini-courses serve as review for the statewide assessment, and most core subject teachers—including the principal and the assistant principal—are responsible for teaching them. In addition to taking enrichment classes, male African-American students are targeted for special assistance in writing and portfolio work because the school's staff

found gender and ethnicity gaps on assessment scores.

Pine County District

Pine District is a large urban district containing a wide variety of schools and school types. There are alternative schools, traditional schools, magnet schools, and "regular" public schools often within the same neighborhood. This complicates the student population demographics of any single school within the district. The phrase "typical school" does not apply very well in Pine County. When we add the socioeconomic interaction with local geography to the mix, the result is an extremely wide range of school characteristics within this single district.

We chose Pine County Elementary and Pine County Middle Schools because they were both in "average" neighborhoods economically and neither had a "special school" designation (they were neither traditional, magnet, alternative, etc.). The two schools chosen turned out to be radically different and to serve very different populations of students despite their common designations and their proximity to one another.

Pine County Elementary School

Pine Elementary serves about 450 students. It is an older building, but seems in good repair and is decorated with colorful student work. Office space is at a premium at the school and many of the councilors and other tertiary staff are housed in temporary offices located behind a curtain in the cafeteria. Several of these persons are only at the school a few days a week and share the space.

One of the principal's major concerns is the number of students at the school. The school is designed to house about 400 students, so they are currently above capacity. In this district, that means that additional students who want to register at the school are bused elsewhere. The school must also maintain a certain ethnic diversity, and it is located in a predominantly white neighborhood. Schools in predominantly black neighborhoods face similar problems, so the schools effectively trade students. Pine must send a certain portion of white students, often from its own neighborhood, to other schools. Pine receives African-American students from other schools and neighborhoods.

Pine County Middle School

While Pine County Elementary is probably as close to a "typical school" as is likely to be found in the county, Pine County Middle is not. Pine Middle is located only about five miles from the elementary, but it is a significant five miles. The neighborhood is still "average" economically for the county, but the dispersion of ethnic groups is considerably greater. This school is also in close enough proximity to several low-income housing projects that it serves a large number of impoverished students.

Pine Middle is also affected by the other middle schools surrounding it. At the middle school level the special school designations are coupled with student entrance criteria, parent involvement requirements, or other requirements that limit the enrollment of certain students. Pine Middle has no such requirements and is hence left admitting many of the students who could not attend elsewhere. This system effectively tracks students into certain schools rather than into

programs within a single school. Academically gifted students have a tendency to attend schools with special academic programs. Less gifted students cannot gain entry.

This creates a variety of challenges for Pine County Middle School. First, the surrounding schools tend to outperform Pine academically. This contributes to a poor reputation for academics at the school, making it more difficult to recruit talented students. Special education students often do not meet the requirements to enter the other schools, making the proportion of special education students at Pine abnormally high. Pine District, in an effort to bolster academic performance in low-performing schools, has created an assistance program for schools with scores of less than 40 on the KIRIS assessment. Pine perpetually receives this assistance, but the assistance comes in the form of an education specialist who requires additional professional development for teachers and scrutinizes practices. The assistance is referred to by teachers as a sanction for performing poorly rather than as an aid for improving student test scores. Students who attend Pine Middle tend to have a high transience rate as well, but it is a different transience problem than encountered at many other schools. Pine has negative population growth as the school year progresses. Students with the option of attending another school often transfer before the year is finished. Pine has reasonably high student population numbers at the beginning of the year requiring them to hire significant numbers of new staff members. As the population dwindles they are forced to scramble to keep teachers.

The staff at Pine Middle falls into one of two categories. They are either "missionaries" or they are "defeated." Many of the teachers see the students as disadvantaged and struggle to provide them with the means of overcoming their backgrounds. They work in this school because it provides them with access to those students who need them most. Others are resigned to the idea that irrespective of their efforts, many if not most of their students have very little chance of success. They focus on behavior and social skills, often to the exclusion of academic content, and simply try to survive the year.

Spruce County District

Spruce is a Western Kentucky county. The economics of the area have a strong agricultural base, but despite the rural setting, the area also has numerous industries. Impoverished students do attend schools in the district, but their numbers are smaller than in most other rural districts participating in the study.

Spruce County Elementary School

Spruce County Elementary is a rural school on a small road. It is in a very quiet setting, surrounded by farm fields. The school population is predominantly Caucasian.

This appears to be a school in transition. One teacher described the path the school had taken over the past few years as moving from very traditional to very reform, and now back toward a more middle ground. One interviewed pair of teachers said they would like to try "specialty teaching" (for example, instead of one teacher teaching four core subjects to a group of students, two teachers might divide the core subjects according to the individual strengths of each teacher.). However, because one teacher was reluctant to try specialty teaching, they had decided

to postpone it until everyone could enter it wholeheartedly. Some fifth-grade teachers are already using a form of specialty teaching; one teacher sends her class next door for science instruction, while the other teacher's class comes to her for social studies instruction.

For the most part, teachers at this school claim that they are unable to predict whether their chances for earning rewards will increase or decrease under the new testing system. This may be directly related to the fact that this school initially had been in reward status and then in decline, only to discover that they had been wrongly classified as being in decline.

Spruce County Middle School

Spruce County Middle School has a student population of more than 700 students. For the most part, each grade has its own separate hallway where its classrooms are located. The school has received several additions since being built about 20 years ago, and it is located a short distance from town.

Technology receives major emphasis at Spruce Middle School. The media center sports a computer lab in one corner, and we were able to observe some students working on an elaborate computer/video production, which was to be entered in a contest. The school has more than 400 computers for its students. The music program has been updated, with each student learning music on Yamaha keyboards complete with headsets.

This school is in the second year of the 10 Sigma program, and nearly every teacher mentions it during interviews. It seems to have made a major positive impact on their school. The principal said that 10 Sigma helps with a district-wide curriculum alignment to national and state standards.

Grades in this school are divided into teams, which are able to establish their own schedules as needed. One seventh-grade team, for example, uses a block schedule routinely, while the other team uses it infrequently.

This school has been through major shifts in the assessment and accountability system. The first accountability cycle found the school in rewards, but the next cycle found it in decline, complete with a state-assigned Distinguished Educator. The principal attributes this wide variation in scores to variations in the groups of students who took the assessment. He said he believed their original high scores resulted from an especially bright group of students who did well and established a high baseline for the school, which the second group of students was unable to maintain. The school is currently back in rewards and seeing progress.

Cedar County District

Cedar County is located in the northern part of Kentucky. Many district employees were educated in Ohio or Indiana. The proximity to other school systems and state education reform initiatives provides the teachers in this district with comparisons that are rare in other parts of Kentucky. There is a wide variety of population dynamics in the county. The northernmost portion of the county might be considered urban, but the majority of the northern half would be classified as suburban. The southern portion of the county has considerable agriculture and could be

considered rural, despite its proximity to a major city.

Cedar County Elementary School

Cedar Elementary is located in the suburbs. The surrounding homes indicate that the area is reasonably affluent. It is an older school building currently undergoing fairly major renovation. Primarily, electrical work is being completed to allow for classroom Internet access. The classrooms are large and not typically crowded.

Cedar Elementary was recently redistricted to account for a newly built elementary school nearby. The population of students attending Cedar dropped substantially. The number of teachers at Cedar was proportionally reduced. An effort was apparently made to keep experienced teachers at Cedar while many of the younger teachers went to the newer facility. The median number of years of experience possessed by a teacher at Cedar is more than 20.

The schedule at the elementary is a blend. Each teacher teaches reading, but the other classes are divided such that teachers can teach toward their strengths. This is the first year for this schedule and according to the principal and teachers it is working better than the previous self-contained system. The students rotate through science, social studies, and writing classes.

Sample student work is posted in the halls and in the classrooms. It is typically of average-to-high quality with a few excellent examples. Decorating the school does not typically include common writing guidelines for open-response questions such as the "four-column method" (for a description of the four-column method, see Thacker, Hoffman, & Koger, 1998). Instead, the writing materials that are posted have been extensively interpreted or modified by teachers.

Cedar County Middle School

Cedar Middle School is less than a mile from the elementary, so geographic features remain mostly constant. It is a large facility serving more than 600 students and is about 12 years old. Like the elementary school, Cedar Middle predominantly serves white students. Very few African-American students attend either school. One of the most striking features about the school is that the administrative offices have a row of windows that overlook the cafeteria. Also, most of the teachers have a shared office area located between two classrooms.

Cedar Middle School has taken the teaming concept to an extreme level. Teachers control most aspects about their teams, including schedules. No two teams seem to be on the same schedule. The only common times are enrichment classes that cross teams and the lunch schedule.

This somewhat unconventional scheduling system, which seems chaotic when described, actually functions smoothly within the school. The lack of uniformity between the teams' schedules is mirrored by a lack of uniform behavior standards from team to team and even from teacher to teacher. In some classes students were required to sit quietly at their desks while doing assignments or listening to lectures. In others, the students were spread from the classroom into the halls working on a variety of projects and lessons. Students seemed to understand each teacher's or team's expectations and their behavior reflected the accepted limits. The students we observed were compliant and generally on-task.

Hickory County District

Hickory County is located in rural central Kentucky. The two schools we visited serve the one small town in the district. The rest of the county's students attend other schools within the district. Hickory Elementary and Middle serve most of the county's African-American population, nearly all of the students who live in low-income housing projects, and a large number of the children of the county's wealthier professionals (doctors, attorneys, bankers, etc.). This makes for a much more diverse student population than is found in other schools within Hickory County.

Hickory County Elementary School

The elementary school has two distinct sections, old and new. The old section is the portion of the school that houses the primary grades. It is indeed old and needs some repairs. The new section holds the fourth and fifth grades. The students only moved into this section of the school this January. The new section of the school was still being landscaped during our visit.

Hickory Elementary has a new principal this year. She has high aspirations for the school but still recognizes some of its problems. Hickory has a very young staff as well. One of the teachers we observed was just completing her internship year and another was waiting to be hired full time in order to begin her internship.

Hickory County Middle School

Hickory Middle School serves the same population of students as the elementary school. The facility is about 5 years old and serves about 400 students. The grades are divided into three distinct halls in the school. Each hall has its own meeting room and work area. Common rooms like the library, cafeteria, industrial arts room, and music room, are located in a section away from the classrooms. There is a great deal of separation between the principal's office and the classrooms.

A running theme discussed by teachers and administrators at both schools in Hickory County was that the school was treated unfairly by the accountability system because of the population it served. The principal at the middle school referred to a segment of the school's population as fourth-generation welfare recipients. The school has more than 50 learning disabled students and about 70% of its students receive free or reduced-price lunch. Those numbers seem daunting when compared to the more rural schools in the county, and may be contributing to the low student expectations observed at these schools.

Poplar County District

Poplar is a rural Western Kentucky county with a population that is undergoing some significant economic changes. The area has been hard hit by cutbacks in the coal mining industry; the loss of jobs has rippled out to the loss of businesses in the community. Because of the changes in economics, the population has become less stable. Families move into and out of the county regularly creating high transience rates for most of the schools within the district.

Poplar County Elementary School

The ethnically diverse student population of Poplar County Elementary School reflects the changing population of its surrounding community. Due to the economic downturn, Poplar Elementary has lost about 100 students since the principal came on board several years ago, and it currently has a student population of slightly more than 400. Redistricting also hurt this school, according to the principal, as it sent some of their better students to other elementary schools in the county. The reduction in school population translates into the probable loss of a staff member next year.

On the other hand, the losses have been somewhat offset by growth in lower SES student population. The establishment of a job-training center nearby has created a pocket of public housing, and those students are bused to Poplar Elementary. The school is also seeing an increase in the number of students whose families are moving from larger cities and bringing "inner city" problems with them. The school is dealing with these challenges in several ways; it now has a child care center for about 100 students and an Extended School Services program in which all students may participate. At the time of our visit, about 22-25% of the student population was minority; 70% of the student population was eligible for the free/reduced lunch program.

Another area of concern is the increase in the number of special education students. The principal says that for every regular classroom teacher she loses, she is able to replace that person with a special education teacher. During one testing cycle, fully 35% of the students at the school had been identified as special education students, with Individualized Education Plans or 504 Plans.

The faculty has spent the last year working on curriculum alignment, both within the building and within the district.

The principal says that she plans on switching her third- and fourth-grade teachers next year. She says the former fourth-grade teachers will be better able to prepare the third-grade students for the assessment, since these teachers have a better idea of what to require of the students. The former third-grade teachers will gain experience, as well, by taking fourth-grade students through an assessment.

Poplar County Middle School

The loss of jobs in the community was also a concern at Poplar County Middle School. As an example of the job loss suffered in this small community, one teacher said that when she first settled in the area, the town had several clothing stores; when she returned after having been gone for several years, there was only one left. Some teachers say the middle class has been sharply reduced in this area. They are concerned about the effect that this change may have on their school's scores. The student population is about 600, but this is expected to drop by 10% next year.

This school is working on curriculum alignment and is in the early stages of the 10 Sigma program. There is a core of teachers who are already involved in the program, and one of these teachers said that they didn't want to rush into things. The school also will look at block scheduling and teaming for the next school year, partly to help eliminate scheduling inequities of some teachers. During our visit we saw several teachers who had to prepare lesson plans for

several subjects and grade levels, while others had to plan only for a single subject at a single grade level. They are also considering moving from a 12-week grading period to a nine-week grading period in order to accommodate more elective classes.

Teachers were dissatisfied with the accountability system. They said they feel it is unfair, particularly when a group of students at the elementary school earns rewards, and then when that same group arrives at the middle school it goes into decline. Finally, that same group ends up in rewards again at the high school.

Cottonwood County District

The Cottonwood district is in the mountainous eastern portion of the state where coal mining and related businesses are the principle sources of income. It is difficult to call the area "rural" in the typical sense of there being farm acreage and large spaces between residences. Valley floors dictate housing; land for schools is carved from hillsides.

Cottonwood County Elementary School

Cottonwood Elementary School is located in the primary town in the county. It contains roughly 425 students, with approximately half on free or reduced price lunches. Fourth and fifth graders receive their instruction in the core subjects of math, science, social studies, and language arts by rotating between four different teachers, each of whom specializes in one of the core topics. This rotation has been in place for about five years and was originated to reduce the amount of time teachers in self-contained classes devoted to developing students' KIRIS portfolios. Students are orderly in their transition between classes and do not appear to be unduly disrupted by the movement. Classrooms, however, felt crowded because of a lack of storage space.

Cottonwood County Middle School

Cottonwood Middle School, also located in the county's primary town, is housed in a rather new, spacious building. Drawing students from deeper in the valleys than Cottonwood Elementary, the numbers of free and reduced price lunch students is somewhat higher than for the elementary school. Discipline, however, does not appear to be a problem. The school has experienced both Distinguished Educators and Highly Skilled Educators over the last few years, signaling its struggle to achieve KERA's expectations. As a result, they are investing extra effort in reading programs and have shifted their schedule for teaching science and social studies. Seventh grade students are enrolled in two science classes. Eighth grade students are enrolled in two mathematics classes and two social studies classes. The school has also started a separate, one-semester, economics class for seventh graders. The middle school principal is aggressive in acquiring school improvement grants. For example, they have acquired funding to construct a biology study area behind to school and have applied for a Comprehensive School Reform Demonstration Grant. In general, teachers are positive in their outlook.

The middle school has also started working with the high school to help solve their reading problem, but have not yet begun to coordinate with the elementary school. The district is also beginning to recognize the need for district-wide coordination of curriculum, and they sense that

teachers would like to have an integrated curriculum that makes instructional expectations clear for all grades.

Locust County District

Locust County is an urban district with some variation of schools and school types. It is not as varied as the Pine County district, but parents have some choices regarding which public school their children attend. There are traditional and magnet schools in Locust County, but the majority of schools are not specially designated.

The two schools we visited in Locust County were within two miles of each other in a reasonably affluent suburban part of the county. One of the aspects that initially drew us to choose these schools, in addition to our normal selection criteria, was the separation of scores between the middle and elementary school. Locust Elementary serves essentially the same population of students as Locust Middle, but the elementary's 1998-99 KIRIS scores were more than twice as high. It wasn't until we contacted the middle school that we discovered that their score is largely reflective of penalties incurred during the last KIRIS test administration.

Locust County Middle School

Locust Middle School is in "KERA Jail," according to one of its teachers. The school is contesting its score report for 1998-99 because of penalties associated with administering the KIRIS tests. Some teachers issued students a copy of a blank four-column method worksheet. The sheet is designed to help students organize their thoughts before answering an open-response question, and while posters depicting the columns are allowed in the rooms, providing students with the worksheets is a violation of testing practice. Students who received the blank worksheet had their scores nullified causing the school to drop well below its goals for the year.

Many Locust Middle School teachers are mistrustful of anything coming from the Department of Education. No substantial analysis of test scores was used in school planning at Locust. Most teachers still don't know what their students' test scores would have been if not for the penalties. Efforts to improve instruction through professional development were secondary to "code-of-ethics training." Locust Middle has been beaten up by the accountability system and they aren't likely to begin healing until their battle over the 1998-99 KIRIS scores is resolved. Despite those issues, however, the classes we observed at Locust Middle were challenging and engaging in comparison to most middle schools participating in the study.

Locust County Middle School

Locust Elementary, on the other hand, is very proud of its test scores. The staff has made marked improvements over the past few years and has received monetary rewards as a result. They are innovative and independent. The principal at Locust Elementary is very proud of the fact that the programs and policies they have implemented over the past few years were the result of school-level decisions.

Locust Elementary is a large school with more than 600 students. The facility is older, but in good repair. Classrooms are large and typically decorated with student work. The staff

represents a wide range of years of experience.

Walnut County District

Walnut County District is located in the coal-mining area of southeastern Kentucky. The high school and middle school are contained in the same relatively new school located at the edge of the small town that is also the county seat. The high school occupies the bottom floor while the middle school occupies the top floor. The two schools operate as totally separate schools with different school colors and team names. These are the only high school and middle school in the district. As is typical of the area, a large number of students receive free or reduced meals—71% at the elementary school and 84% at the middle school. The town itself appears bustling. This may account for the relatively lower percentage of free or reduced meals for the elementary school. Based on observations and interviews, the DAC came across as being one of the most active and involved of all the administrators encountered in our visits. This was her first year in the district office, which had suffered from being understaffed. The administrators, principals, and teachers were friendly and appeared to talk openly after overcoming their initial nervousness or anxiety about the data collection. However, some of the teachers at the elementary school appeared to retain their apprehension about the interviews and observations. Students were generally polite and orderly and were inquisitive about our visit.

Walnut County Elementary School

Walnut County Elementary School is a K-6 school with about 400 students. Several years ago, all elementary schools in the county lost their seventh and eighth grades when a new middle school was created. The fact that sixth grades have remained at the elementary schools instead of becoming part of the middle school has created some problems, which administrators at both the elementary and middle schools we visited recognize. The principal at Walnut Elementary, for example, has offered to send the sixth-grade teachers to the middle school faculty meetings and professional development offerings, and has even suggested that the middle school principal observe them on a regular basis. Apparently little has resulted from these overtures, although the elementary principal says that some teachers do work with middle school teachers on an "individual basis." Apart from that problem, however, there appears to be little concern with curriculum alignment issues at the elementary school.

Walnut Elementary School operates on a self-contained basis for fourth and fifth grades; sixth-grade students appear to rotate out of their classrooms for one or two core subjects only.

The administrators say that this school initially moved from traditional classroom instruction (students seated in rows in very quiet classrooms) to instruction based more on reform practices (small group work, less emphasis on remaining quiet, less use of textbook as the sole source of knowledge). They also focused on elaborate school-wide thematic units that were popular with students and teachers, but they were disappointed when their KIRIS scores failed to go up. Their Distinguished Educator told them that they had failed to concentrate on the Core Content. Now, the administrators say, they are returning to an emphasis on the Core Content and the "basics."

Walnut County Middle School

Walnut County Middle School has a school population of slightly more than 400 students in its seventh and eighth grades. The school's principal is also concerned that sixth-grade classes have remained with the county's elementary schools; he says he believes that middle school scores would improve if the sixth grades were to become part of the middle school. However, he does not want to push this issue. The communities want to keep the sixth-grade students at the elementary school level and in the local community.

The principal says that the middle school is in the early stages of aligning curriculum with the high school and elementary schools, with the major push for the alignment process to take place in the summer. Teachers at this school generally seem to teach both grade levels of a particular subject, and they seem more aware of curriculum alignment and what it is designed to do. Science teachers expressed particular concern over this issue, especially as it related to sixth grade. The middle school is also reluctant to give up teaching the thematic units, however.

The number of special education students has nearly doubled since the school was built several years ago. The principal says these students are mainstreamed for at least 80% of their instruction. School officials are thinking about ability grouping students next year to better meet the needs of their high- and low-ability students.

According to one teacher, the school has had to cut back on the number of teachers over the past few years and has been unable to replace them. Those who remain have to be more flexible in their teaching assignments, especially those with the K-8 certification, the teacher added.

Structure of the Study

This portion of the study is divided into four sections, the first three of which directly relate to the research questions. The first section addresses the issue of instructional practice. The second section addresses communication issues, and the third section addresses confidence and perception issues. The fourth section addresses other important issues, such as special education, that arose during the site visits.

Instructional Practice

What changes in classroom instruction have taken place as a result of the change in testing from KIRIS to CATS?

A major objective of our research for this school year was to determine whether the transition from KIRIS to CATS has caused a ripple or a wave. In general, the conclusion is that in comparison to KIRIS the introduction of CATS is causing no more than a ripple. When we directly asked educators if CATS would stimulate changes in teaching and learning, the resounding response was "No." The prevalence of this response was based on their beliefs that CATS does not represent a significant departure from KIRIS. As we heard from numerous teachers, "rumor has it that CATS is KIRIS with a new name." Whenever teachers elaborated on that answer, however, several different reasons emerged.

Changes that resulted from KIRIS continue to be implemented

The first variation of the "no change" theme is that KIRIS had a big impact on instruction

and those changes are expected to continue. A sixth-grade science teacher from Pine County commented that "Ten years ago there was no guidance and no one to help. Now there are resource persons, and performance standards that eliminate the easy way out, that is, always using worksheets and puzzles." A seventh-grade social studies teacher from the same school added that while "good teaching has been constant and included a mix of content and application of content, KIRIS emphasized that focus." An eighth-grade teacher in another part of the state echoed that her instruction would not change because she has always used hands-on work, projects, and class presentations. A Cedar Elementary teacher said, "I've already made needed changes for KIRIS." In short, "KIRIS had a big effect and I assume that it will continue," quipped an elementary teacher from Poplar Elementary.

Tests should not change instruction

Another position underlying teachers' responses that CATS would not influence their teaching was that a test per se should not change instruction. Sentiments included comments such as "You need to teach students, not the test," and "It shouldn't. Good instruction should be the same regardless of testing method. There is no perfect test." Other teachers more explicitly referred to instruction needing to be guided by curriculum content. For example, an eighth-grade social studies teachers said that the CATS test "will not impact and that teachers should not teach to the test. I have a set curriculum; for example, I teach the Constitution because the students need to know it." Teachers from two of the larger districts indicated that their instruction is driven by district curriculum guidelines more than by the test. Finally, using a more negative tone about the test, an eighth-grade social studies teacher said that he did not agree with the test's emphasis on open-response because it limits coverage of history facts which are needed in order to answer open-response questions. In other words, this teacher was looking for more balance in instruction than he thought was represented by his understanding of the test's design.

Test is unrealistic for student population

The third variation on the theme that the test will not impact instruction came from teachers at the middle school we visited in Pine County. These teachers face an atypical population of students, and for them, the CATS test is simply not highly relevant to their day-to-day concerns. In addition to its population of mainstreamed students with behavioral disabilities (BD) and learning disabilities (LD), this school has a large population of students in self-contained BD and LD classrooms. Teachers from these classes told us:

- "The relevance of the test to LD students is problematic. LD students are so low functioning that the assessment does not impact them. The questions are not realistic even though several have alternative portfolios and the rest have accommodations."
- "Topics and performance standards are too advanced for BD students. Our focus is in a different place."

The following comment typifies teachers for the remainder of the school. "Any consistent test can have an impact and in theory, all students can learn. In reality we already have had over 40% student turnover of students on our family team this year. The relevance of education and student

motivation is such a problem that adaptation to the test is a minor consideration."

These comments, however, should not be taken out of context. The teachers at Pine Middle School were operating in arguably the most difficult situation that the research team has experienced in its three years of school visits. Furthermore, these teachers were engaged in a variety of practices commonly associated with adapting to the open-response format of the old test. Because of their consistently low scores and not unexpectedly low gains,[1] the school has been inundated with state and district support personnel who have stimulated reform efforts. From our perspective, the pressure for instructional change stemming from KERA has been at least as great in this school as other schools around Kentucky, but that pressure pales in comparison to the daily pressure created by the student population.[2]

"Wait and see"

The final variation on the theme that the introduction of CATS would not impact teachers' instruction had a strong temporal aspect. That is, teachers said they did not yet know enough about the CATS test for it to influence their instruction, but that instruction "may shift emphasis after seeing the test, " and that "learning more about the test will influence instruction." For example, they will "probably try to clue in on what to look for in the test because their test-taking strategies are not real good right now." "We'll teach toward CATS as more is learned about it."

Closely related to the "No, but later" theme are the teachers who responded in the affirmative—that CATS would influence instruction. Two classes of "yes" responses were discerned, however. The first is a "Yes, unfortunately/with reservations." For example, the following comments all seem to express the same idea:

- "Whenever there is a test program, there is pressure to teach the test, not necessarily by local people, but by the whole education system. This year the school is working on blending content and process more – last year was more writing process emphasis. Teachers have the attitude that they should teach the students what they need to know, which will help them in high school versus teaching what is on the test."
- "There is much pressure to teach the test so I hope that the test will cover the appropriate content – i.e., what is in the respective curriculum goals."
- "After they see the content of the test and learn about the rotation of topics, the test will probably influence instruction. The goal is to get all students to master the Core Content, but not all students are alike and not all can learn at the same level."
- "Yes, what is being tested will be emphasized versus just teaching the Core Content and Program of Studies.

[1] Medsker (1998) has shown that gains in test scores are low for Kentucky schools with high rates of student transience.

[2] For example, a teacher told us about asking a particular student why he smelled so bad and the middle school student said that he had spent the night in jail. As another example we left one of our classroom visits early after a police officer arrived in response to the teacher pressing her "panic" button – a communication device to alert the office that she needed help. The students were not violent in this case, but they were uncontrollable in their disregard for any attempts at instruction. Office personnel said that this teacher very seldom called the office. We later asked the teacher if our presence might have triggered the students' reaction. She indicated that the problem had started before we arrived.

Notice in these comments a skepticism about whether the test will match the Core Content for Assessment, which is the document intended to define the domain of test content.

The second class of "yes" responses is a "yes, and that's OK" attitude. The supporting details, however, indicate that these views do not necessarily stem from a motivation simply to increase test scores. For example, one teacher elaborated that "teaching to the concept of the system can open minds and create critical thinking." We also heard a comment indicating that a teacher was using the content coverage targets set by preliminary test blueprint workshops last summer to modify his/her instructional coverage.

With regard to instructional changes, our basic finding is that, in general, KIRIS-initiated changes in instructional practices are being perpetuated. Changes in practices this year are the result of ongoing change initiated under KIRIS and the lack of clear understanding on the part of teachers as to what the CATS test will actually look like. Whether instruction will be modified in the future appears to depend on how teachers react after CATS has been in place.

What types of instructional practices are being used in classroom instruction?

This question is not as easy to answer as it may initially appear, because several factors work together in influencing classroom instructional practices. In the first part of this section, we consider what practices are being observed, and then follow with the effect of teacher content knowledge, available resources, and other factors on instructional practices.

Reform versus traditional instructional activities

Over the last several years, national organizations such as the National Council of Teachers of Mathematics, the National Council for the Social Studies, and the National Research Council have issued new teaching and assessment standards for their members. These standards, the earliest of which was written in 1989 (National Council of Teachers of Mathematics), state that students should be able to use knowledge rather than just recall facts and dates. The NCTM standards, for example, call for the emphasis of practices such as extended and open-ended problem solving and manipulative use, while de-emphasizing basic computation skills and drilling.

In Kentucky, this emphasis on reform practices shows up in documents available to teachers, such as *Transformations: Kentucky's Curriculum Framework* (KDE, 1995), which give teachers examples of reform methods to use in their classrooms. Students are expected to use higher order thinking skills such as "compare," "analyze," "research," and "develop" in their daily work, and these skills are also found in the higher levels of Bloom's Taxonomy (Bloom, Ed., 1956).

During our initial visit to 20 Kentucky middle schools two years ago, there was evidence that teachers were beginning to switch from more traditional teaching methods, such as extensive lecturing, to more reform methods, such as group learning and use of manipulatives. Use of these newer practices was positively correlated with gains in KIRIS scores (Hoffman, Harris, Koger, & Thacker, 1997; Harris, Hoffman, Koger, & Thacker, 1998). In general, however, students in the 1997 observations tended to work on worksheets, view material written on overhead projectors, and use the text as the primary classroom resource, even though the classrooms were stocked with

a variety of materials such as computers, TVs, and other subject-related books. Teachers also spent more time instructing students on the proper method of doing a four-column method piece or an open response question. Some classes would work through these together, and the teacher would spend much of the class period directing their efforts. Now, however, open-response questions appear to be more familiar to students. There did not appear to be as much emphasis on working through them as a class, although they were routinely assigned as homework and given as part of classroom assessments. In one instance, where a teacher was observed discussing her students' open response test performance, the discussion was a detailed diagnosis of both the content knowledge and the process steps required to achieve a high quality answer.

The impact of teacher content knowledge on instructional practices

The problem that several teachers seem to be having is attempting challenging techniques and failing to follow them through to completion. This was especially evident in fourth-grade science classes that we observed. One teacher divided her class into small groups of three or four students each, and then gave each group an experiment to complete. She was highly organized, having previously set up each experiment so as not to waste time, and she made sure that students understood that they were to use scientific inquiry (form and test hypothesis, record and analyze data, and state conclusion) while doing the experiments. However, the group that we observed was unable to understand the "why" behind their experiment (hard-boiled egg being forced through the narrow neck of a bottle due to differences in air pressure). They offered such explanations as "the air pressure is greater in the bottle" or that the "heat (from burning match in bottle) caused it to go partway down." These observations were offered without clarification or correction from the teacher or questioning to lead the students to discover/correct their erroneous thinking.

A fourth-grade science teacher at another school had her class, working in groups of two, conduct an experiment proving that evaporation cools. However, she jumped from this simple concept to determining the relative humidity in different areas of the school and finally, to having her students design miniature "coolers" that worked through evaporation. The teacher added to the confusion in this lesson when she told the students to think of "coolers" they had at home. Unfortunately, the example that came readily to mind was the common Styrofoam cooler, which works by insulation, not evaporation. The students appeared to have no idea what they were doing or why they were doing it when they designed their coolers. In fact, the teacher had to direct their efforts rather substantially with hints and clues. Students with a better understanding of the concept likely would have been able to design their coolers more independently.

Thus, lessons that began with great promise—using many reform methods such as hands-on work and cooperative learning—ended up being less effective than they could have been because of insufficient content knowledge.

Teachers who successfully engage their students in genuine interactive learning face a challenge that is not intuitive at first glance. The successful teacher must deal with the curiosity and momentum she has created. Students who are coerced, cajoled, bribed, enticed, strong-armed, or even tricked into learning are often not satisfied with the explanations and descriptions of some idea or concept that is included in their text book. In short, kids ask good questions. It is up to the teacher to determine whether these questions add or detract from the class. In the best case, the questions fuel the learning of the student who asked it, other students, and perhaps even the teacher.

In the worst case, the question reveals some degree of lack of content knowledge on the part of the teacher and either stops the lesson, allows misinformation to be distributed as genuine, or simply allows the lesson to move tangentially to some unrelated topic.

At least two teacher factors interact to determine how successful student engagement will be handled. The teacher must know what information to teach and the teacher must know how to accomplish the teaching. We have all heard the adage that "knowing what to teach does not make one a teacher," but it is equally true that "not knowing what to teach prevents one from becoming a teacher." Ideally we would like for all teachers to be experts in their fields of study as well as highly skilled instructors. Unfortunately, this is not always the case. The teachers we observed during this study exhibited a wide range of both teaching skill and content expertise.

A middle school teacher from Hickory County used a Socratic questioning methodology to teach students about the layers of the Earth's interior. She kept the students excited and engaged throughout the lesson but the lesson lasted only about 15 minutes. She moved on to a discussion of geology and then to a discussion about determining the age of rocks and then to the methodology for mapping the ocean floor. The transition between topics was sporadic and disjointed, but through skillful presentation and enthusiastic speech the teacher was able to keep the students' attention. The students, however, seemed to know that the teacher's content expertise was weak. They used the questioning format of the class to jump around within the topics and to extend the discussion beyond the teacher's comfort zone. The first time a student asked a question for which the teacher had no answer she suggested that the student "look that up and tell us about it tomorrow." The students seemed not to take the suggestion seriously. Students were told to "look that up for us" at least five more times during the one-hour class period. The teacher would also occasionally ask students to predict the answers of their own questions. The students' responses were consistently incorrect, but they were told, "I don't really know, but that sounds like a good theory." Students were allowed to expound greatly on incorrect themes with no questions or corrections to guide them back to a more reasonable discussion or even toward the day's lesson agenda. This is not to say that the teacher was unprepared for the lesson. On the contrary, she seemed to know the information provided in the textbook nearly by rote. Her expertise simply did not extend much beyond the level of the middle school science text.

This example was not isolated during our school visits. We made no effort to count content mistakes made by teachers during the course of this study, but they were often noted in observers' logs. We did note the methodology by which teachers interacted with their students. Most teachers seem to be attempting to use reform practices suggested by KERA documents. If those practices succeed in having students take a more active role in their own education, then the quality of interaction between teachers and students must necessarily improve. Teaching methodology may be improving in Kentucky, but improvements in student performance could depend on improved teacher content knowledge. This problem is not isolated to Kentucky; indeed, there is a national trend to attempt to improve teacher-content knowledge, which has been determined to be deficient across the country (Archer, 1999; Galluzzo, 1999; Ingersoll, 1999).

A positive example of artful combination of content knowledge and instructional skill was observed in a middle school science class. The lesson was on the earth's rotation, tilt, and revolution around the sun. After some introductory demonstrations, using students as role-players,

the teacher pulled out two goose-necked lamps and two thermometers, placed them on the table, and asked the students to design an experiment to show how the tilt of the earth creates the seasons. In our interview after the lesson, this veteran teacher said that just a few years ago he would have just told them what to do. Instead, he patiently solicited idea after idea, redirecting mistaken suggestions with the simple phrase, "Does anyone have a different or better idea?" With few exceptions, the students eagerly participated in the design of the experiment and were prepared to carry it out in the next day's class.

A final concern about teacher content knowledge is the difficulty that some teachers have in taking graduate courses in their content areas. For example, we talked with a teacher who loved teaching in his content area. He wanted to continue teaching in that content area and possibly teach that subject at a small college or university in the future. This teacher was working on a master's degree in counseling although he had no desire to become a school counselor. The degree in counseling was the only degree he could realistically obtain to meet the Kentucky master's degree requirement. He told us that although he would prefer a master's in his content area, he was unable to pursue that degree without taking off from teaching. Graduate classes in his content area were not offered locally and the almost two-hour, one-way commute to the nearest university was considered unreasonable. Additionally, graduate classes in his content area were not offered other than during the normal school day. A teacher at a different school also complained that taking content area courses required a one-way commute of at least 90 minutes if she could find the courses to take.

There are numerous areas within Kentucky that require long commutes to a location offering graduate courses. Although there are many branch campuses that offer graduate courses, many options available for teachers at branch campuses do not include programs in a content area.

Another teacher who lived near a state university also complained that he was unable to take courses toward a graduate degree in his content area. As a result, he was getting his graduate degree in education administration. The content area courses offered were either not graduate-level courses, were offered during normal school hours, or were extended courses taught on weekends throughout the semester.

Although this was not an area that we were investigating in our research, this is an issue that should be investigated. With a call for teachers, especially at the high school and middle school levels, to be better grounded in the content area they are teaching, efforts need to be made to make graduate-level programs available for them in specific content areas.

The impact of classroom materials on instructional practices

In nearly all the schools we visited this year, an abundance of materials was present. In addition to textbooks, there were videodisks, computers with Internet access, software programs, workbooks, student planners, science equipment, and other resources in evidence. Perhaps even more impressive was the nature of these items. They were specifically designed to teach students to access higher order thinking skills and were often aligned with national content standards. Since Kentucky's Core Content also mirrors national standards, the result is teachers armed with more and better-quality teaching tools with which to perform their jobs.

Even the textbooks are oriented toward reform practices. They still have the obligatory chapter reviews full of vocabulary words, multiple choice and short answer questions, but they also have supplementary information allowing for activities, using manipulatives, varying learning styles, accommodations for slower and faster learners, portfolio entries, and discovery learning. Teachers have told us in the recent past that they had been instructed to move away from reliance on the textbook, and indeed they had done so (Hoffman, Harris, Koger, & Thacker, 1997) but textbook manufacturers may be giving them a powerful incentive to move back.

Putting these tools into practice, however, requires considerable skill on the part of the teacher. Teaching, and therefore learning, depends on implementation. Videodisks were used in at least four of the observed classes in three different schools. The results of each demonstrate the importance of implementation irrespective of the teaching tool.

In the first, the videodisk was used as a poster board. Pictures were placed on the television screen and referred to by the teacher. The lesson was nearly exclusively teacher centered. Students sat at desks and heard a lecture. Interaction between students did not occur and interaction between the teacher and student was sparse.

In another classroom, the videodisk was used to provide background to the teacher's interchange with the students. While the videodisk showed scenes of modes of transportation, the instructor began introducing the topic of fossil fuels. The connection, however, was stretched.

In the third, the videodisk was used more effectively, but the lesson contained within the vignettes shown on the disk was difficult to understand due to some teacher editing. A Cedar Middle School teacher was conducting a lesson about AIDS, but was opposed to any mention of condoms or other contraceptives. Those references on the disk were omitted. The result was a choppy lesson containing video of a few persons diagnosed HIV-positive who gave testimonials about tolerance and the perils of promiscuity. The questions designed to generate discussion, contained in the guidebook that accompanied the videodisk, were written down and answered individually by the students. The students were effectively lectured to by the videodisk instead of the teacher. Interaction was minimal and the separate portions of the lesson were never pieced together into a cohesive framework for thinking about AIDS.

The fourth video lesson was conducted masterfully. The Elm County Middle School teacher used vignettes from the disk to illustrate the differences between physical and behavioral adaptation. The disk provided examples of each and the teacher used a very Socratic discussion format to check for understanding among the students. The students were energetic and engaged during the lesson and asked pertinent questions of the teacher and each other. The teacher demonstrated a good deal of content knowledge and was comfortable both in answering students' questions and in telling the students how to discover the answers on their own. The lesson was opened with an exciting hook to gather the students' attention and closed with final thoughts about how behavioral and physical adaptations might be thought of together and how to discern one from the other. Students' responses to questions showed that they had generalized the concepts associated with biological adaptation and could differentiate within those concepts. Higher order thinking was required for the discussion and demonstrated by the students who were participating.

These examples illustrate the necessity of skillful implementation. Providing teachers with

materials, irrespective of the quantity or quality of those materials, falls far short of ensuring quality learning. The potential, however, for improved application in the future is enormous.

The impact of self-contained classrooms on instructional practices

Instructional practices also may be influenced by the continued implementation of self-contained classrooms in elementary schools. This practice, once the standard for elementary schools around the country, appears to be taking a downturn in the schools we visited. Replacing them are more specialized classrooms in which elementary teachers may be responsible for only one or two subjects.

The main advantage to this system is that students are likely to receive instruction from teachers who are more comfortable and knowledgeable about the subject. Having this specialized teaching (some schools call it "departmentalized instruction" or "skilled area teaching") prevents a subject from becoming "an area that gets shortened at the end of the day," said one teacher of fourth- and fifth-grade science.

There is some evidence that content knowledge in self-contained classrooms remains a problem in some districts, particularly with science instruction. However, several other factors may contribute to the problem, as well. These factors may overlap, making a "clean" analysis impossible, given the limited nature of our school visits. These factors can include:

- The limited number of science courses that teacher candidates were required to take, particularly those teachers who earned K-8 certificates.

- Effect of teacher gender on science instruction. Traditionally, elementary teachers are women, and women in the past were less likely to volunteer to take higher level science courses both at high school and university levels (Henke, Geis, & Giambattista, 1996; Smith, 1995). This may contribute to less content knowledge about science and a lower comfort level when teaching it.

- Sixth grades that are still located in elementary schools. Given that Kentucky's science assessment is administered in the seventh grade, this situation means that middle schools are held accountable for students who have been in the school for only a few months. Two middle/junior high schools in this study had to deal with this situation, and teachers and administrators recognized that it was a problem for their schools. In both districts, at least some sixth-grade classes were still completely self-contained; others were using variations of departmentalization, which ranged from mostly self-contained to mostly departmentalized. Seventh-grade science teachers at these two schools were particularly concerned about the quality of their students' previous science instruction. One teacher said she assumes that her students "don't know anything when they get here." Although most of her school's feeder elementary schools are departmentalized, several are not, even at the sixth-grade level, she noted. These students are at an additional "disadvantage" because they have not had to learn basic organizational skills such as bringing the right book to class. They are also less able to handle responsibility, the teacher said, because too much was done for them as sixth graders. A second seventh-grade science teacher expressed concern that elementary teachers in the fifth and sixth grades were not emphasizing science as much as they were supposed to, since

students had already been assessed in science in the fourth grade and would not be assessed again until the seventh grade. In both districts, cross-building communication that could have reduced these problems had been attempted, but had been difficult to maintain.

- Incomplete curriculum alignment. Note that this issue is closely related to the previous factor; in one district, the problem created by having sixth grade elementary classes was compounded by the fact that the district has only this year begun the curriculum alignment process, with most of the alignment to take place in Summer 1999. This lack of an aligned curriculum makes it difficult for teachers to know what to cover—or what previous teachers should have covered. A third seventh-grade science teacher, for example, said she hoped that sixth-grade students had covered what they were supposed to in life science; she has de-emphasized that topic this year in order to give more time to physical science, which had been a problem area on their KIRIS scores.

The influence of off-assessment grade practices on instructional practices

Students were not tested in every subject of the KIRIS test each year. They took certain sections during the 4th grade and others during the 5th, for example. They took tests in those same subjects in the 7th and 8th grades and again in the 10th and 11th grades. CATS is similarly arranged, with the exception of the norm-referenced test to be given in the 3rd, 6th, and 9th grades.

Schools have taken various interesting approaches aimed at maximizing their test scores and the year that specific content tests are given has garnered much attention. Some schools have chosen to adopt a block scheduling system to emphasize tested subjects, giving them more time than "off-assessment" subjects during the testing year. The amount of extra time varies from school to school, with some schools choosing to eliminate off-assessment year subjects altogether. These schools have the equivalent of two science classes and no social studies in the seventh grade and two social studies classes and no science in the eighth grade. Other blocking regimens are less extreme.

Other schools have chosen more subtle means of coordinating subject area teaching with the state's testing schedule. Off-assessment classes are redefined to more closely match the testing regimen. For instance, in a couple of the middle schools we visited, eighth grade science classes emphasized more of the practical living core content, which is tested in grade 8. Students still attended science class, but the topics they learned about came from practical living/vocational studies curriculum guidelines. They studied health issues, AIDS or other science topics that might help them on the eighth-grade section of the test.

In yet other middle schools, the focus has shifted to more general writing and reading skills. Since the KIRIS test, and presumably the Kentucky Core Content Tests, required so much reading and writing from the students, several schools have boosted requirements for reading and writing within or instead of content requirements. Many have chosen to offer Content Area Reading (CAR) classes in off-assessment grades. This measure gives the students exposure to the content information, but stresses the importance of reading and writing skills. As a more extreme example, some schools had shifted their block schedules to reflect the perceived importance of reading and writing. Those schools often required students to attend classes focused on reading, writing, working on portfolio entries, or similar exercises for more than one half of their time at

the school each day. The other content teachers were forced to split the remaining time between them.

Academic freedom and its relation to instructional practices

When the *Core Content for Assessment* document (KDE, 1996) was released several years into the testing program, its authors emphasized that the material found in the document was to be considered as only one part of a school's total curriculum. The local curriculum, for example, "should address national standards, identified community topics, and other content/skills/performances which the state assessment cannot address," according to the document's authors. Core Content guidelines were to be included in the local curriculum, but they were "not intended to be curriculum standards" nor to "reflect a state curriculum," nor to limit that (local) curriculum," the authors stated (KDE, 1996, ix-x).

Despite what the authors specified about how the document was to be used, the Core Content has become what many Kentucky schools are using as their curriculum standards. The Core Content document has become the "bible" in many schools; many teachers said they base their teaching on its guidelines and are hard pressed to cover what they are supposed to before the assessment. This gives them little time to devote to things outside of the Core Content guidelines. Schools are altering their curriculum to focus on the Core Content document, effectively expanding their curriculum while simultaneously limiting the topics and/or techniques from which they can choose to teach. For example, administrators at a small-town elementary school described how their school had gotten involved in elaborate, whole-school thematic units on topics such as the Oregon Trail and the Middle Ages. Staff members were disappointed when their social studies scores failed to go up, but their Distinguished Educator told them that they were not teaching the Core Content. They have resolved to stick more closely to it in the future, in the hopes that their scores will rise. They seem to regret not being able to present these units and they believe the units were valuable in their own right. In fact, when we asked the administrators what one school-wide project or program had had the biggest influence on students and teachers, they mentioned those same theme-based units. A middle school teacher in another district also discussed the influence that the Core Content has on her teaching. She said she has been "forced to re-examine the Core Content…to be more aware of what has to be taught across the state rather than just what I want to teach."

A few teachers, on the other hand, ignored the Core Content guidelines on what to teach or when to teach it if they thought it was important for their students. One middle school social studies teacher, for example, involved his students in thematic units such as lifestyles of Native Americans. This unit, for seventh-grade students, took about six weeks to complete, and it included a culminating activity that involved the creation of an Indian village on the school grounds. Local elementary students toured the village, where they saw displays of corn being ground, deer being skinned, and the hides being tanned. His eighth-grade students took part in a videotaped "documentary" of the Holocaust, complete with swastikas hanging from the school building and video shots of glowing furnaces. A seventh-grade social studies teacher in another district was teaching about World War I, even though the Core Content suggests that eighth grade students study United States History through the Reconstruction.

What level of work is expected from students?

From its beginning, KERA has emphasized that students must be able to use critical thinking skills such as analysis, evaluation, and synthesis (KDE, 1995). However, if teachers require little beyond fill-in-the-blank worksheets from their students, then those students may not get the opportunity to practice and master those higher level skills. The result is ineffective teaching and learning.

In this study, the work expected from students varied greatly from teacher to teacher as well as from school to school. At Oak Elementary School, for example, proficient work was used as a starting point from which the class worked on making the proficient work better. Several teachers at Oak disclosed that they were expected to have students produce distinguished work. One science teacher had 3 distinguished students and 16 proficient students. As a result, they were working with their students so that the students could both recognize and produce distinguished work. This school had a very diverse student population with about 40% of its students eligible for free/reduced meals. This was the only school that was observed using specific examples of what would be graded as proficient. Several other schools used scoring rubrics that explained the various levels but did not provide any concrete examples.

At Hickory Elementary School, one of the classes had produced newspapers of events leading up to the American Revolutionary War. The class had worked in groups to produce newspapers that included news stories, pictures, advertisements, etc. that could have appeared in newspapers in Boston during this period. Stories included the Boston Tea Party, Boston Massacre, Stamp Act, etc. Most stories were very brief with few details. Some of the stories were not stories at all but consisted of phrases put together. Although the concept of the newspaper required the groups to be creative and attempt to experience the time period, the quality accepted by the teacher for this project was far below what was displayed at Oak Elementary.

A middle school teacher from Cottonwood gave her class the entire period to answer an open-response question comparing two forms of government. This open-response question would be worth half of the student's weekly grade. According to the teacher, this class contained a mixture of top students and low-performing students. The previous class period had included developing a table of comparisons between the two forms of government. That class was basically prewriting in preparation for the open-response question. The students were instructed to write a rough draft of the response, have another student review and make suggestions, and then revise the paper prior to turning it in at the end of the class period. Students were limited to the front of one page of paper for their response. Although students were grouped in pairs or small groups throughout the room, no student was observed following the teacher's instructions as far as having another student review and comment on their paper in order to make revisions. Instead, students turned in the papers to the teacher and picked up "filler" work to occupy the remainder of the period. The teacher accepted the students' work without comment.

In several schools, teachers provided class time for students to do homework. Discussions with teachers and administrators at these schools indicated that many students would not be able to do any work at home for reasons beyond the students' control (normally some kind of a dysfunctional home environment). In several classes we observed, almost 50% of the class was allocated for students to read assignments or complete worksheets, which would typically be done

as homework. Contributing to the use of classroom time for homework completion may be the policy, in at least a couple of districts, of purchasing classroom sets of textbooks rather than purchasing enough copies for each student to have his or her own. Elm Middle School, for example, used some of the money they saved on textbooks to purchase a laser disc player for the school, however they assured us that there are enough copies of textbooks for students to sign out as needed.

The problem of low expectations exists even along with exemplary lessons, such as the Elm County Middle videodisk lesson on physical and behavioral adaptations previously discussed. These students were assigned seatwork after their discussion. Unfortunately, the seatwork did not match the preceding discussion in terms of its complexity or required depth of understanding. The assignment involved coloring a map and predicting the migration paths of butterflies and could be described as remedial. It was also only tenuously related to the discussion. The directions were vague, and interpreting them was more challenging than the content of the assignment. Students spent a great deal of time deciphering the instructions and coloring their maps, but almost no time was devoted to making predictions about the migration paths. No time at all was spent linking the map exercise to the discussion of adaptation. The resulting student work was of predictably low quality and added little in terms of reinforcement, extension, or even assessment of student understanding. A quick examination of the student work posted on the bulletin board in the room showed that this type of student work was not atypical. Student-designed food webs were displayed that were of low quality and were often grossly incorrect. The posted work was neither corrected nor graded. Other examples of student work displayed in the room were similar.

What influence has the addition of multiple-choice questions had on instructional practice?

When Kentucky Department of Education officials issued guidelines for the development of a new assessment system, they included instructions for the development and inclusion of multiple-choice questions as well as for the open-response questions, which had been strongly identified as part of the old KIRIS assessment system. In their instructions, KDE officials specified that multiple-choice questions were to be designed that would assess content knowledge which could be considered basic skills (KDE, 1998a). Most of these questions would assess content knowledge using lower level thinking skills (e.g., knowledge and comprehension); a smaller percentage of questions would assess content knowledge using higher order skills (e.g., application, analysis, synthesis).

Teachers in this study tended to believe that multiple-choice questions measure content knowledge, while open-response questions measure process skills. They generally were enthusiastic about the inclusion of multiple-choice questions to the testing format. In many schools teachers said that because other tests that students would likely be required to take were multiple-choice format, the Kentucky Core Content Tests would provide valuable practice. Tests such as the ACT, SAT, and GRE were mentioned as examples. Other teachers said they believe multiple-choice questions are “less subjective” than are open-response questions, and that they can be graded more accurately than open-response questions. Several teachers listed an advantage of multiple-choice questions as being able to test more content, instead of hoping that they had covered the few topics assessed by the open-response format. Finally, several teachers said that they believe multiple-choice questions will aid those students who are not strong writers but who

know the material.

Evidence suggests that teachers are taking seriously this adjustment to the testing format. Some teachers, for example, mentioned that they have begun teaching their students multiple-choice test-taking skills as a part of their curriculum, while others said that specific practice on the multiple-choice format is necessary because in previous years they had moved to one that was much more dependent on open-response questions. Some schools and districts have purchased multiple-choice practice tests ("Test Ready") for classroom use, as well. Teachers also reported that they were including more multiple-choice questions on their tests, and the principal at Elm Middle School said that a Scantron® test-grading machine has gotten increased use this year after several years of very little use. Given that teachers still appear to rely heavily on publisher-created tests, most of which have sizeable portions of multiple-choice questions, it is doubtful that many teachers got completely away from this testing method in the intervening years.

What influence has CATS had on teacher professional development?

Teacher professional development has improved, according to teachers, since the implementation of KERA (Thacker, Koger & Koger, 1998). Teachers have a great deal more choice regarding what types of training to attend and a great deal more funding is available to fund professional development (McDiarmid et al. 1997). Teachers and schools have begun to use professional development as a means of addressing weaknesses identified by student test scores.

At the time of this study, however, teachers only rarely reported having attended training specifically regarding CATS or the Kentucky Core Content Tests. The training that teachers did report having attended fell into two categories. A few teachers reported that their school district had held training sessions about the changes in the calculation procedures to determine a school's accountability category. From descriptions of these sessions it seems that the primary focus of the meetings was the percentage that each portion of the test would count. Little or no attention was given to what the changes meant in terms of instruction or school governance.

The other type of professional development teachers reported involved the "code of testing ethics" that each teacher who administered KIRIS and will administer the Kentucky Core Content Tests is expected to follow. Teachers in several districts were evidently given a refresher course on the "dos and don'ts" of test administration. Teachers who attended this professional development session reported that it was a repeat of similar training related to KIRIS administration, with no new rules associated with CATS.

Even though teachers report that professional development has gotten better, dissenting opinions were found at the district level. In one district, a representative lamented the changed role of the district office in choosing teacher professional development. She claimed that many teachers lacked the expertise to choose professional development that would most benefit the school system. "We trust the very same teachers who put us in the bottom fifth percentile before KERA to decide how best to improve our schools. The district is forced to advertise and encourage getting teachers to attend training that would benefit their classrooms. We can't mandate in-service training any more, but we probably should." Her opinions were largely isolated and directly at odds with comments from the principal of an elementary school in the same district, but both parties agreed that professional development was a major focus of their efforts to improve

classroom instruction.

Communication

What have teachers heard about CATS?

When teachers were asked what they had heard about the new CATS program, their responses ranged from a predominant, "Not very much," to a much more rare "Quite a lot." When they told us exactly what they had heard, their comments ranged from vague notions about the intent of the CATS program to specific detailed accounts about the status of the program and how it differs from KIRIS. Some teachers recounted what can only be categorized as misconceptions about CATS and the Kentucky Core Content Tests.

Most teachers explained that they had heard very little about the new CATS program. Several teachers who listed many of the specifics regarding the new testing program prefaced their response by saying that they should probably know more. The things that teachers had heard about CATS is less interesting than the strategies they chose to use to explain their knowledge to researchers. Two basic strategies were particularly revealing concerning the focus schools are placing on CATS. Teachers either explained the differences between KIRIS and CATS or they explained the similarities. The teachers choosing to explain the differences focused on the accountability system. Those who used the similarities in their explanations focused on the intent of the testing system. For instance, one teacher explained "Now the test should be less subjective due to the multiple-choice component." Another explained that "Since there is going to be a test and since it would still cover the Core Content, I haven't really given the format changes much thought." In any event the consensus seemed to be that the test itself was simply a renamed KIRIS. The real changes were about what was to be done with the scores from the test after it had been given.

A small number of teachers had serious misconceptions about the new testing system. One explained that he had already taken the CATS test 15 years ago during his own schooling. Several other teachers who were new to Kentucky schools were unfamiliar with CATS or the previous accountability system. Most of these new recruits to Kentucky were also new to teaching, giving them an especially large amount of information to internalize in a very short time. A teacher from Pine Middle School who had been educated in Georgia related that her teacher education program had been closely tied to Georgia's testing system and curriculum guidelines. She has been using those guidelines in Kentucky and reports that they served her well as she began her teaching career. New teachers educated in Kentucky did not volunteer information to suggest so close a link between the states public education programs and their teacher preparation studies.

How have teachers learned about CATS?

CATS has received a great deal of attention in newspapers, on television, on the radio, and in numerous publications such as the *Kentucky Teacher*. However, when asked, teachers most often cited the school where they taught as their primary source of information about the new testing system. "I try and read everything they put in my box," was a fairly common response to researchers' questions about the origin of their knowledge about the new system.

Teachers' informal conversations were the next most frequently cited source of information about CATS. Since not all teachers attend all the same professional development, many have learned to rely on each other to distribute valuable information to the rest of the school. The degree to which information is diluted or altered by the time it reaches its final destination is unknown, but the misconceptions shared by teachers within a few of the participating schools, suggests that the system is not yet perfect. There is still the danger of perpetuating rumor or compounding misunderstandings.

Some of the schools also held faculty meetings dedicated to learning about the new testing system. In some cases, the DAC was invited to come to the school and lead the faculty training. More often, however, a representative from the school attended training led by the DAC and then brought that training back to the school.

A representative from the district office from each of the 10 districts visited was interviewed. During those interviews we discussed the role of the district in preparing schools for the upcoming state assessment. Not surprisingly, the district is a primary communication channel for getting information about testing to the school level. What might be surprising was the degree of variability regarding how this occurred from district to district.

In both of our two urban districts the decision of whom to interview was complicated. These districts have a great many more district personnel than the rural districts and the responsibilities of the district toward the schools is split into smaller pieces in those districts. Persons whose job description seemed best to fit the "curriculum and assessment" category were selected. In the rural districts that usually meant the District Assessment Coordinator (DAC). In the urban districts there were offices devoted to curriculum and assessment independent of the DAC.

This difference of size and complexity plays out as two distinct communication problems, one for the urban districts and one for the rural districts. In the rural districts the DAC is called upon to perform nearly all tasks associated with the state assessment, from coordinating transport of test books to schools to fielding questions from teachers and administrators. They often plan and conduct teacher professional development related to assessment. They also help schools disaggregate the data once it is received and are usually involved in helping the schools develop their plans for improvement. When it is taken into account that most of the rural DACs have taken the job in addition to some other function at the district office, the problem becomes clear. Even if the DAC is well trained and well organized, the amount of responsibility placed on her is daunting. The volume of effort and the number of schools within the district can result in gaps in communication filtering down from the state level to the teacher level. Schools are forced to seek their answers on their own or to wait their turn for the attention of the DAC.

Urban districts have a very different problem. Theirs is a difficulty of coordination. Having a district staff that numbers more than 100 creates the situation where, if one is not terribly careful, certain sections of the district office are performing functions that other sections are at best unfamiliar with and at worst counter to their efforts. A good example from Pine District is the production of the district curriculum guidelines. These elaborate guidelines are available on the Internet and on CD for all schools in the district and were evidently quite expensive to produce. Pine Elementary seems to be using them nearly exclusively as their curriculum guide. Irrespective

of the quality or origin of these documents, they have assumed a role in the curriculum planning for the district. They may have taken the place of the state-produced Implementation Guide for the Program of Studies in this district. This example represents only one of the problems associated with a district this size.

Another major issue is the overlapping information sent to the schools. Teachers often recounted being inundated with materials from the district, the state, or from the school. They claimed that they got so much material that it was simply impossible to read and sort through all of it. Many of them hadn't read our letter describing this study because they assumed that it was "more of the same" from the district. Pine District schools typically had an impressive array of materials from the district posted on the walls. There were laminated posters depicting Bloom's taxonomy, the Four-Column Method, and others on the walls of nearly every classroom. Each had the district insignia printed on it, suggesting a massive production effort. The advantages of a well-staffed, well-funded district office are numerous, but in many cases the very things that give them their advantages work against them. Schools and teachers in these districts often have the same concerns and questions as their counterparts in rural districts. Clear and concise information flow from the state level to the school level depends on the effectiveness of the district offices. Whether due to a lack of sufficient staff and assistance or an overabundance of it, this flow of communication remains in a logjam for many of the schools we visited.

Although KDE maintains an Internet site, very few district personnel and even fewer school-level personnel mentioned the Internet as a source of information. Most schools had Internet access and e-mail available to teachers, many with access directly from the classroom. These resources are either not being utilized or the teachers and district persons who participated in this study did not consider the Internet a resource worth mentioning.

On the other hand, an effective district communication program exists in the Walnut district, where the DAC produced newsletters for distribution to the schools about testing. She was a former assistant principal at the middle school and had been in the district office only since October. She listed numerous activities she and her assistant had instituted, and she gave the impression that her recent stint in the school had served her well. She seemed to have a good idea of what the schools needed to know. Someone more removed from the school may not have been as effective.

Are there other communications issues related to CATS?

Team teaching

Other aspects of communication relate more specifically to communication between schools and between grades in the same school. At Oak Junior High, for example, two 7th-grade science teachers with rooms right across the hall from each other popped in and out of each other's rooms during class changing time and talked at lunch about what had worked well for them. One teacher had been doing a lab activity, and she told her colleague about changes that would improve the lab when the other teacher had her students perform the same activity. They were not teaching the same thing at the same time, but they seemed to have a good understanding of where the other was, and they mentioned that they shared assessments. On the other hand, there was little

opportunity for informal communication by the 8th grade science teachers because their classrooms were very far apart from each other.

Oak illustrates examples of both strong teaming effort by teachers as well as teaming in name only. Team concept schools are very common in Kentucky, especially at the middle school level. Elementary schools have only recently begun moving away from the self-contained classroom structure and have had less teaming opportunity. Many middle schools refer to their teams as families, and they effectively divide the middle school into several smaller schools within the same building. Each family has a set of core teachers who teach within their own family of students. A large school of more than 500 students might have families consisting of fewer than 100.

Another major advantage of the team concept is that teachers on a given team often share a planning period. Teachers were observed planning individual educational plans (IEPs) for special education students, preparing for classes, meeting with parents, or performing other duties that related to their teams. Not all staff members from all schools participated in the team meetings, but the opportunity for within-school communication has greatly increased since KERA was implemented.

Curriculum alignment within and across districts

An atypical example of cross-district communication was found at Oak Independent. In discussions with teachers and administrators there, the problem of transfer students was consistently brought up. Evidently a segment of the population moves frequently during the school year, and as a result, children in this more transient segment transfer from school to school several times during the school year. This movement is between schools within the district as well as to schools in neighboring districts. This movement is especially prevalent for students with parents in the lower SES segment of the population. Teachers and administrators understand that this situation is very disruptive to the student's education. A solution being considered to aid these students is having a common curriculum within grades in all schools within the district and the neighboring systems. This would mean that students transferring between schools would walk into a class that was covering almost the same material that was being covered in the class the student left. For example, an eighth grader transferring from School A would be studying the War of 1812 in social studies. When the student enters the eighth grade class at School B that social studies class would also be studying the War of 1812. As it was, one seventh grade class at Oak Junior High School was studying World War I and the other seventh grade class was just beginning to study the Greeks and the Romans. Of course, World War I is probably outside the seventh-grade curriculum. Other seventh-grade classes throughout Kentucky were at approximately the same point—Greeks and Romans—as the seventh-grade teacher at Oak Junior High.

The common curriculum is probably more of a problem in the non-social studies subjects since most social studies classes follow a common timeline in their instruction. Science may be especially problematic because of the extremes we have seen throughout Kentucky.

School Reliance on Documentation

Many schools participating in the study seemed to be searching for a single ultimate guide

to improving accountability scores. Instead of relying on their own interpretation of the various documents that have been released by KDE, their respective districts, and independent publishers, they discover one particular document that is attractive to them and treat as if it were "step-by-step instructions." They then build policies and programs around the document. They use it to plan professional development, alter curriculum, reallocate funding, and for other purposes.

The merits of using a single guidebook are obvious. One need only consult a single source to resolve dilemmas. It is easier for the school to maintain uniform codes of conduct, grading, etc. if everyone gets their guidelines from the same source. It streamlines many school processes and might make issues like curriculum alignment more easily resolved. Instead of deciding who teaches what during the fourth grade, one only needs to consult the handbook.

For most of the participating schools, there was no single handbook. The closest document they had to a guide was the *Core Content*, and they used a variety of other documents to supplement and clarify the information contained within it. However, the Elm Middle School staff has adopted the *Implementation Guide to the Core Content for Assessment* as a prescriptive teaching manual. They are using the sample lessons in the guide to plan their curriculum. If a particular concept is not included in a sample lesson, it is assumed that the concept will not be included on the assessment. These concepts are the first to be cut from the curriculum. A science teacher explained that he was very pleased with the specificity of the Implementation Guide because it made his curriculum more manageable. "I don't have to teach the cell anymore," he said. The implementation guide lists a variety of methods teachers might adopt to teach specific concepts from the *Core Content*. Elm has chosen one of those methods and is rallying all the teachers in the school around it.

Another extreme example was found at Pine Elementary. Pine has constructed its own curriculum guide and distributed it to all the schools in the district. Each school in the district received a paper copy and a CD, and they have access to the document on the Internet. The production of the guide represents a tremendous undertaking on the part of the district. It is unknown how well the information contained in the Pine document matches the *Core Content*, national standards, or other materials, but Pine Elementary is using it as the primary reference for curriculum at the school. Content is specified for each subject at each grade level in the guide and a follow-up document with sample lessons is planned for the near future.

At Locust Elementary teachers received Test-Ready, a test preparation guide that is supposed to produce higher test scores. This guide is not specific for Kentucky's curriculum or testing system, but is a commercially available test preparation program. Locust teachers had only received it the morning of our visit, but many had already mistakenly assumed that it had come from KDE and was endorsed as a testing preparation tool for the Kentucky Core Content Tests.

Practitioners and Their Confidence/Perceptions About CATS

Does CATS represent an improvement over KIRIS?

A predominant complaint about the KIRIS system was that teachers, administrators, parents, legislators, and other concerned parties had lost confidence that the system allowed for the fair classification of schools. Since this classification was used to justify rewards and

sanctions in the high stakes accountability system, that lack of confidence amounted to a debilitating problem for the system.

Although the Kentucky Core Content Test portion of CATS had not been give at the time of this study, teachers were asked to estimate how confident they were that the Kentucky Core Content Test would be a fairer tool for administering rewards or assistance than KIRIS. Although most of the teachers were aware of the substantive changes to the testing system, very few expressed a great deal of confidence that the changes represented a real improvement. Instead many claimed that they could not know until they had actually seen the test and received results. Others hedged their optimism by saying, "It can't be worse than KIRIS." Neither of these answers amounts to a rousing endorsement of the Kentucky Core Content Tests, but these two opinions encompass the overwhelming majority of the participating teachers.

The next most common response can be categorized as a cautious hopefulness. A few teachers expressed that they would need to wait and see, but they followed up by saying that they were optimistic. A few teachers claimed that they were not confident in the new accountability system at all. Most of these teachers also expressed philosophical opposition to any accountability system based on student test scores, however.

"Apples to oranges"

KIRIS, with its intention of changing education, caused many teachers to be resentful when told that they would have to change their teaching methods to accommodate the new educational system. KIRIS and its accountability system, which required that schools meet interim goals, came under a barrage of criticism from teachers. By far, the most common complaint heard in our three years of visits to Kentucky schools is that rewards and sanctions should not be based on scores from different student cohorts. These teachers declare that "It's like comparing apples to oranges!" Instead, they believe that students should be measured against themselves in some form of longitudinal study, and they also believe that will result in a fairer assessment system. Many teachers attribute their schools' initial high baseline scores to an unusually gifted student cohort; they said this high score could not be maintained by the more ordinary cohorts that followed, and their schools were then labeled as being "in decline." In a related concern, teachers blamed variation in student cohorts for wildly varying scores—in decline one cycle and in rewards the next, or vice versa.

Many teachers in our site visits expressed the idea that KIRIS had been unfair because it was "too subjective," "ambiguous," or "vague." Several teachers mentioned specifically that the heavy writing emphasis of KIRIS was unfair to students for whom writing was a weaker area. Generally, teachers said they thought the inclusion of multiple-choice questions to the CATS test would be an improvement to test fairness.

One teacher apparently either failed to understand or did not agree that, under KIRIS, schools were to be measured against their own past performance, instead of being measured against the performance of other schools. This teacher thought it unfair that other schools with scores "15 points below our baseline" were receiving rewards while his/her school wasn't eligible for rewards.

Teachers still express complaints about KIRIS scoring. Some teachers, for example, said that their top students scored poorly on KIRIS, while students who did poorly in their classroom did well on KIRIS. A middle school social studies teacher complained that some students at his school scored a "98" on the social studies portion of the test, but were classified as "proficient" rather than as "distinguished." He also claimed that there were no "distinguished" social studies students in the entire state, and that this was unfair to students and teachers.

Some teachers believe that the way in which special education students are assessed is unfair. A middle school teacher, for example, said that special education students have received "proficients," while his top students get "novice" ratings. Another middle school science teacher (who had previously been a special education teacher) said he did not think that special education students should take the same test as non-special education students. He also said the KIRIS test failed to recognize the real achievements that some special education students made. These students may have raised their performance from low "novice" to high "novice," he noted, but that the KIRIS test only recognized movement from "novice" to "apprentice." An elementary teacher who was in her second year of teaching said that remedial students or those with IEPs may need extra help on the test, but not qualify for accommodations. And finally, an elementary principal said that her school had lower ability students who did better than higher ability students on the on-demand writing assessment. This may have been because the higher ability students were not specific enough in their writing and tended to write too much, she theorized.

Test manipulation

Several teachers from different schools talked about various forms of test manipulation they said they had heard about: improving their chances of rewards by deliberately scoring low on baselines, "playing the testing game" or "working the system" for rewards, and "teaching to the test."

Concerns about setting baselines

Cedar County Middle School has had both a modicum of success and a few setbacks since KERA began. The school originally set a low baseline when KIRIS began and then immediately posted enough improvement to qualify for rewards. They were held up as an exemplar by the district, but went back into decline during the third testing cycle. Several teachers recounted the fluctuation in their previous scores as an indicator that the KIRIS system was flawed and inaccurate. "We went from having other teachers come in to learn what we were doing to having a Distinguished Educator tell us what to do in under a year. We didn't change anything that we did. We didn't teach any worse or better. We just got our turn, I guess," explained one teacher.

Not everyone at the school remembers it that way, however. Another teacher at the same school explained that the school's original low baseline was the result of inaction by the school. "Once the baseline was set, the whole staff really buckled down and worked hard with the kids. We got rewards and a lot of the staff slacked back off. Surprise, surprise, we went into decline." She also claimed that the school hadn't yet recovered. "They know it's another baseline year, so I don't expect our scores to improve much this cycle."

During our interview with the Cedar County principal, he mentioned that it might be

advantageous for a school to set a really low score during the first CATS accountability cycle. He was aware of the funds available for schools in decline and that the Highly Skilled Educator (the replacements for the Distinguished Educators) could be refused, effectively eliminating the possibility of imposed sanctions. He also said, "Of course we'll all try our best on the test. Kids here just try and do well on all their assignments and I think the staff has some pride at stake in the new test scores."

Setting low baseline scores was also mentioned by a Locust County middle school teacher, who said that he was unsure of CATS' ability to help the state administer rewards or assistance more fairly. He mentioned that some schools may have learned to play the testing game better than other schools by deliberately setting a low baseline score in the first year.

The testing game

This concern differs from the previous one in that the complaints that teachers expressed are more generalized. Several teachers said they did not have faith in the reward system because of alleged cheating, which they described in general terms. As one middle school teacher said, "People will work the system whenever there is money involved…there are too many ways to play the right games to get rewards." A teacher from another school declared that s/he was "against rewards" because there was too much cheating by participants in the testing system.

Teaching to the test

A related concern is that of "teaching to the test," which teachers thought was a way of manipulating the test, versus teaching what they thought their students needed. A middle school teacher said, "Teachers here have the attitude of teaching what they think is important—not teaching what's on the test." He explained that he saw that as a positive aspect of the school's teachers. And, finally, an elementary school teacher commented that teaching to the test is less effective: "If teachers are teaching to the test, students are more apt to forget—if teaching to students, they will remember in the long haul. That's what I'm after."

Other teachers were concerned about certain testing procedures, which, although allowable, may be unfair to certain groups. An example is the fact that the series of KIRIS tests, and now Kentucky Core Content Tests, must be given in a certain order. One middle school science teacher complained that this consistently penalizes science and social studies, which are given at the end of the testing period, when students are "burned out" with testing. Concerns about the testing of special education students are addressed in a later section. One middle school teacher also complained about specific questions on the KIRIS test, which s/he said had come from areas outside the Core Content and thus should not have been included on the test.

Will CATS affect student preparation for middle and high school?

When teachers were asked about the effect that CATS would have on their students' preparation for high school (middle school for elementary teachers), the overwhelming response amounted to, "Not any." There are a variety of possible reasons they might have responded in this manner, however. First, they may indeed have concluded that the accountability and testing system had no effect on their students' preparedness for the next higher-grade level. Second, they may

have misunderstood the question. The majority of the teachers interviewed used "CATS" to refer to the Kentucky Core Content Tests. If they thought the question referred to the Kentucky Core Content Tests in isolation from accountability formulas and other components of CATS they may have interpreted the question too narrowly. They may have considered only the actual act of completing a Kentucky Core Content Test booklet. If so, it is not surprising that they were skeptical that CATS could influence their students' preparedness. Or they may simply have omitted consideration of the effects of the accountability system on the school and the implications of those changes for the students who attend it.

The few teachers who did respond that CATS would affect their students' preparedness had only modest hopes. They commonly claimed that the test would allow the students to become familiar with multiple-choice format tests, and that this familiarity might translate into better preparation for other multiple choice tests such as the ACT, SAT, GRE, ASVAB or other examinations. A few other teachers said that they would need to wait and see what effect the system had on their own or other teachers within the school's instruction. If the system required changes, then the students' preparation might be altered as well. These teachers did not venture a guess as to whether the likely change would be positive or negative.

Is the reward system fairer under CATS?

Scores from the KIRIS test translated into either monetary rewards or school sanctions. CATS maintained the reward system, but altered the method of giving rewards by issuing the money as school funds instead of allowing teachers to distribute it as bonuses. We heard about schemes for distributing the funds ranging from dividing monies evenly among teachers, administrative staff, and non-certified staff (including lunchroom staff and bus drivers) to designs that included only certified teachers. The common opinion seems to be that allowing teachers to administer rewards caused more trouble than it was worth.

Many of the teachers we spoke to candidly disagreed with the idea of any kind of reward system for schools. Typical statements from teachers included:

- "I don't think anyone teaches better or harder because they might get a check from the state,"
- "The rewards just caused a lot of hurt feelings,"
- "I wish this school had never gotten rewards at all,"
- "I never thought the rewards were fair. You know as well as I do that there are teachers in every school that do a good job, just like there are teachers at every school who try and beat the kids to the parking lot every day,"
- "Teachers do a good job for the satisfaction of knowing their kids are learning something, that they make a difference. I doubt many of them teach for the chance to get a reward check."

These statements indicate that a large number of teachers have philosophical reservations about school rewards. These opinions are not scarce and are not limited to schools that rarely, if ever, received rewards.

A second, slightly different opinion concerning the rewards involves the emotional stress of administering the money. One district administrator explained that while being a reward school brought a certain amount of pride in their accomplishments, teachers were not "the type of

professionals for which rewards provide incentive." She further explained that teachers at one school felt so guilty about receiving rewards when teachers at other schools had not, that they spent the bonus checks almost exclusively on school supplies and equipment. "They knew that those other teachers had worked just as hard and received nothing." The administration of funds by teachers evidently caused considerable stress due to feelings of guilt on the part of the teachers responsible for the distribution.

The history of the teaching profession may help explain why a reward system, a proven strategy in other businesses, was not the incentive it was meant to be. The teaching profession has never been particularly competitive. Prospective teachers might compete for a position in a school when they are initially hired, but that is generally the full extent to which one teacher is formally compared to another. Teaching remains, for the most part, an isolated profession (Meyer & Rowan, 1977). Once teachers close their classroom doors, the only likely scrutiny their teaching will need to endure for the year is one or two observations by the principal and, if they are teaching in an assessment grade, the state accountability test.

What effect will putting CATS scores on student transcripts have?

A portion of the requirements included in House Bill 53 that established CATS included an evaluation of the reliability of student Kentucky Core Content Test scores for the purpose of including the test scores on students' transcripts. Scores would then follow the students from elementary to middle to high school, and presumably beyond. House Bill 53 does not include specific information regarding expected or recommended uses of these scores.

Teachers and principals were asked their opinions about the effects of placing scores on students' individual records. Most reacted positively to the idea of including any form of student-level accountability in the system. However, they were also quick to point out that they expected very little change in student motivation at the elementary and middle school levels. Many said that their students were simply too young to understand the implications of transcripts. A few of the middle school teachers from the urban districts said that if scores on transcripts were used by local high schools as an entrance criterion, their students would pay more attention to the test. Elementary school teachers were nearly uniform in their opinion that it would make no difference. The principal from Hickory Elementary suggested a positive implication of placing scores on elementary records. She posited that the test might be taken more seriously by some of the parents in the district and that they might then do more to encourage their children to do their best. On the other hand, she also said that most of the students whose parents would take notice already did their best on the test, as well as any other work they were assigned. Most of the teachers interviewed said that the addition of scores to transcripts might be taken much more seriously at the high school level. They claimed that this would be especially true if the state colleges and universities took the scores into account for determining admittance or eligibility for scholarships.

Many teachers and administrators also said that while they applauded student-level accountability, placing student scores on transcripts might be missing the portion of the student population with which they are most concerned. The common opinion is that students who care about things like grades and transcripts are the same students who are already doing their best on the state accountability test. The students who do not care about transcripts and grades are the students who require some form of external motivation to give their best effort on the test.

How useful were the 1998 KIRIS score reports?

The 1998-99 school year might be considered the "lame-duck" year for KIRIS. As the accountability system changed to CATS, schools were awaiting their final KIRIS score report. The scores from the 1997-98 academic year still held considerable significance for the schools in that they were used to determine which schools would receive rewards and the amount of those rewards. Schools, however, have been trained to use the score reports for a variety of other reasons that might be considered questionable given the change in accountability systems.

How could score reports be improved?

Schools used KIRIS scores in the past as a diagnostic tool for determining their strengths and weaknesses. They have traditionally spent a great deal of effort analyzing the data provided in the report and many have even paid outside consultants to perform additional analyses for them. They then use the data to plan for teacher professional development, curriculum changes, scheduling changes, allocation of funds, and other aspects of school change that they write up in their Consolidated Plan. The Consolidated Plan replaced the Transformation Plan, which required all schools to author an outline of their plans for improvement and the data that support those plans. Once schools had made their plans and carried out their initiatives, the most obvious choice to ascertain if their choices were successful was the KIRIS reports from the next year. Despite a time lag problem (Thacker, Koger, & Koger, 1998), schools used the KIRIS reports as a diagnosis, addressed their problems, and then used the KIRIS reports again as an evaluation of the success of the devised policies and initiatives. This year, however, there will be no new KIRIS report from which to gauge the success of policies and programs. The preparations the school made for the accountability test will impact Kentucky Core Content Tests, not KIRIS. The best hope that schools have is that their preparations for KIRIS will translate to similar results on Kentucky Core Content Tests, but if the change in testing formats is considered significant at the school, the utility of the KIRIS reports is questionable.

Researchers asked school personnel about the usefulness of the 1997-98 KIRIS score reports. They followed with a query about the utility of that report compared to previous reports. In most instances, teachers and principals claimed that their use of the KIRIS reports had not changed. The new CATS system is considered sufficiently similar to KIRIS for the same strategies to be expected to work. In terms of the overall utility of the report, it is still touted as the primary source of data for school planning. The only improvements that were suggested for reporting under the new system included some added analysis so that outside consultants would not need to be hired and to decrease the time lag between when the students take the test and when the report is released.

The expected similarity between KIRIS and CATS in terms of planning policies and interpretation of student achievement brings up a more difficult problem within the accountability system. Schools expect to be able to use CATS scores to indicate and demonstrate growth from 1997-98 to 1998-99, and they expect to be able to continue to monitor student growth. Practitioners have not internalized the significance of the break between KIRIS and CATS. Spontaneous reference to the "interim" accountability period was essentially non-existent.

The reasons for the lingering ideas that CATS will be comparable to KIRIS become obvious after speaking with the principals of the participating schools. The Pine Middle principal explained that he expected scores to improve dramatically this year at his school due to the anticipated addition of points for different levels of novice and apprentice category students, where the vast majority of his students fall. The principals at Elm Middle and Locust Elementary both expect their science scores to improve this year. Elm Middle recently had an influx of energetic talented young science teachers and Locust Elementary activated a special science resource teacher to provide exciting and complex laboratory experiences for students. These principals, and indeed most of the others, expect to measure growth. They expect the proportion of students at their schools who score in the novice category to drop due to the school's policies and programs.

While the State Board of Education, Kentucky's National Technical Advisory Panel on Assessment and Accountability (created by HB 53), KDE, and CATS contractors have thoroughly discussed the extreme difficulties of comparing 1999 scores to 1998 KIRIS scores, the field has no knowledge or recognition of the problem. Expectations that schools can assess their growth during the transition are due in part to the fact that the achievement level category names will remain unchanged after the switch from KIRIS to CATS. Student will still be novice, apprentice, proficient, or distinguished. New descriptors for these categories had not been released at the time of this study, nor were any plans to alter the category descriptions in evidence. If the proportion of students in the novice category changes, the schools seem very likely to assume that it is because of improvements in their instructional methods. The retention of the achievement level category names makes growth attributions very likely. In addition, the *Kentucky Teacher* (Fishback, 1999), a KDE publication, implies that schools' ability to interpret change will be uninterrupted.

Other perceptions which arose during visits

Teachers in assessment grades and increased stress

In the early years of KIRIS, teachers in the original assessment grades (Grades 4, 8, and 11 or 12) found themselves very much "under the gun." It was the performance of their students that was being assessed, and this, in turn, reflected back onto them as an assessment of their teaching. Kelley and Protsik (1996), for example, brought up the difficulties of being a teacher in an assessment year with the following comment by a teacher participant: "The bottom line for me is, yeah, we try to share the responsibility of the assessment, but when those test scores come in, teachers say 'How did the fourth grade teachers do?'" (pg. 43-44)

Now, however, assessments have been spread out among more grades, lessening the focus on any one grade. For example, KIRIS assessments in 1998 were given in the 4^{th}, 5^{th}, 7^{th}, 8^{th}, and 11^{th} grades, and portfolios were required in the 12^{th} grade. Under CATS, every grade from the 3^{rd} through the 12^{th} grades will be assessed, either by the Kentucky Core Content Tests, portfolios, writing prompts, or norm-referenced tests.

The burden on assessment grade teachers also has been lessened as communication between grades within a school and between schools has improved. This has occurred through the curriculum alignment process. This alignment process assigns responsibility for specific topics to

certain grades, as well as preventing gaps and duplication in coverage. While some districts are farther along in the curriculum alignment process, nearly all districts we visited said they were working on some aspect of it.

Even with these improvements, teachers in assessed grades still report more stress than those who are not teaching assessed grades. The Poplar Elementary principal, for example, said she was going to switch her third- and fourth-grade teachers next year. She reported that her fourth-grade teachers were becoming burned out after several years of bearing the assessment burden. When those teachers are moved to the third grade, they would already know what students need to know and be able to do upon entering the fourth grade and would be better able to prepare the third-grade students for fourth grade. When third-grade teachers are moved to the fourth grade, they in turn would learn more about the assessment itself as well as what fourth-grade students are expected to accomplish.

Seventh-grade science teachers also appear to be more stressed than other teachers, particularly if sixth grade students are still at elementary schools. Communication is not routinely occurring between schools, so these teachers are often unsure as to what students have been taught prior to the seventh grade. They feel that they alone are responsible for preparing their students for the assessment.

Other issues

Special education concerns

Although the original plans of this research did not target special education specifically, we would be remiss if we did not dedicate a section of this paper to the many concerns of the special education teachers and others who we interviewed. These individuals have concerns about the relationship between the testing and accountability system and special education students that cut very quickly to the heart of what it means to educate "all" students. Their students are required to take the state accountability test with accommodations that are determined from a list specified in each student's Individual Education Plan (IEP). Likely accommodations range from large-print test booklets to scribes who write the student's answer as it is dictated. As special education teachers see it, the concern is with test validity once accommodations have been implemented.

At first glance, one might think that special education teachers were simply lobbying for exemption from the accountability system. This was not the case, however. At least a couple of teachers knew the adage that "What gets tested, gets taught," and they recognized the significance of including special education students in the state testing program. These teachers want to be included in the system and they see the doors that might be opened for their students because of that inclusion. One teacher explained that since the beginning of state testing she had been included in more of the "regular staff meetings" that concerned curriculum and other issues and that she had altered her own teaching to approximate the topics covered in the Core Content. She said that she still needed help in certain content areas, but that she felt more a part of the school than before. She is working toward garnering more communication and assistance from the school's "regular" content teachers.

The problems with testing special needs students under the current system are diametric

opposites. On the one hand, many special needs students taking the test have little or no access to the questions regardless of any accommodations, making it impossible for them to demonstrate their level of competency. On the other hand, some teachers are questioning the use of accommodations that make it difficult to determine whether the test is measuring anything about their academic achievement.

Accessibility to the test

As mentioned previously, some teachers were concerned that special education students taking the regular KIRIS test, even with accommodations, were not being recognized for the real progress they made. The KIRIS test, for example, only recognized one level of novice performance and was not sensitive to the growth of some special education students who probably would never advance beyond novice performance. Certain schools, with relatively high percentages of special education students, seem particularly hard hit by this problem. Teachers at Pine Middle School, for example, reported that about half of their students read below the third-grade level. For these students, no amount of reading the question, paraphrasing, or even scribing is going to allow them access to the test. They simply do not have the necessary vocabulary and skill to understand what the questions are asking. To categorize all these students into the novice category disregards the possible achievements that they and their teachers might have accomplished during the year. Teachers at Pine want to be part of the accountability system, but they also want some assurance that the test will allow their students to show progress in a way that takes into account where the students began and which is fair for teachers and students. In schools like Pine, where the population of special needs students is inordinately large, this problem takes on a significance that might be masked in other schools around the state. Greater attention to the treatment of special needs students in the accountability formula and in the testing system in general will be required before this issue is resolved to the satisfaction of Pine's staff.

Validity of test accommodations

With Kentucky's emphasis on the belief that all students—even those with special needs—can learn at higher levels, the assessment of its special education students plays an important role in accountability issues. As mentioned before, those students who use accommodations listed in their IEPs are entitled to use them on the regular state tests.

Teachers at several schools, however, expressed concerns regarding the validity and appropriateness of some accommodations. They related stories about how special needs students outscored the rest of the student population. One teacher, for example, said that accommodations unfairly raised the scores of some special education students. "Because of accommodations, there are lots of proficient reading scores when students cannot read—the test is read to them," the teacher noted. The implication in these stories was strongly suggested that those scores are reflective of the accommodations rather than of the students.

Fluctuating Economic Conditions and Their Effects on Student Population

Redistricting can be defined in two ways: Formal redistricting, which results when the board of education formally alters a school's serviced area, and informal redistricting, in which

the school population is altered due to outside influences.

Two elementary schools, one in an independent city district and one in a small-town county district, expressed similar concerns over informal redistricting affecting their school population. Both areas had been hit with job layoffs and business reorganizations, which school staffs feared would translate into the loss of some of their more economically stable student population. These two schools also faced other situations that contributed to informal redistricting:

- School officials at Oak Elementary School said that the completion of a county elementary school would likely reduce the number of county students who currently pay tuition to attend Oak Elementary. They predicted that lower-income students from within their district would replace these higher-income students.

- Officials in Poplar Elementary School said their school population was transitioning from that of the typical small-town to more of an "inner city" due to the creation of a nearby job corps center.

Both schools predicted that their scores would fall as a result of this informal redistricting.

Penalties for Violating Test Administration Procedures

Several Locust Middle School teachers violated KIRIS test administration policy during the 1997-98 academic year. They admit the violation, although they said that they did not know that their actions represented a violation at the time. They passed out a blank 4-Column Method worksheet for students to use during testing. The same worksheet is posted on the walls of every classroom in the school, but posting the 4-Column Method is not a violation while handing out copies of it is. In addition, several of the teachers responsible for the violation said that their state-appointed Distinguished Educator instructed them to hand out copies of the worksheet. The scores of those students who were given the worksheet were reduced to zero as a penalty for the violation and the school's overall index score dropped dramatically.

Not surprisingly, the teachers at Locust were outraged by this turn of events. They are currently contesting their scores. They have also developed an animosity for KDE and for the accountability system that cannot help but affect their efforts toward school reform. The principal at Locust told us that he was still waiting for a score report that might be used for school planning. Many of the teachers had similar claims. Even more divisive is the fact that not all teachers participated in handing out the worksheet, and so not all teachers' students were penalized. Those teachers who were not penalized said that their students had "done pretty well this year." "It's a shame that we won't ever know just how well we would have done without all of this," said one particularly concerned teacher. Another teacher became so upset during our interview that she had to leave before we were finished.

The effects of posting a low score, for whatever reason, can be devastating for a school, but for Locust they were particularly severe. First, Locust is located in a relatively wealthy urban area. It competes with other similar schools for students. Posting a low score and receiving the "in

decline" or "in crisis" label seriously damages the school's reputation. Add to this the stigma regarding the circumstances of the penalty and it becomes clear that teachers in this particular school have good reason to be demoralized. They assume that part of their community believes that they cheated on the state accountability examination and the rest simply assume that they are incompetent. And as if that weren't enough, this situation occurred under the leadership, if not the direct instruction, of a Distinguished Educator assigned to help the school improve.

The severity of the unspecified sanctions that Locust incurred lead one to wonder about the "fit" between the severity of the punishment and that of the crime. Clearly, severe penalties are designed to ensure that test administration is taken seriously and performed uniformly throughout Kentucky. However, the system, of which these penalties are a part, is designed to foster improved instruction and more effective education of Kentucky's students. These two goals come into conflict in the instance of Locust Middle School. The importance given to the penalties have supplanted the rest of the system and affected all aspects of the school. The delays in resolving the contested scores serve only to fuel further animosity and to postpone the school's recovery.

Conclusions and Recommendations

The transition year between KIRIS and CATS is a year marked by caution and skepticism on the part of teachers. They are waiting to pass judgment on a new accountability system that they fear is simply a renamed rehash of the old one. There are two ways of looking at this conclusion. The good news is that most teachers have not abandoned the instructional practices that led them toward more successful teaching under the KIRIS system. They are still using the reform practices that they were trained to use and there has been very little "back to basics" movement on the part of the schools participating in this study.

The bad news is that teachers are not convinced that the system represents a significant change. They therefore have little confidence in the new system. The more enthusiastic of the teachers interviewed claimed that "It's got to be better than KIRIS," but they had little to say about the quality of CATS, choosing instead to focus on the lack of quality of KIRIS. Perhaps even more problematic is the idea that scores from CATS will be misinterpreted due to this lack of confidence that the system has changed. Schools expect to be able to gauge change from the scores they are scheduled to receive in September. They expect to be able to attribute changes in scores to the effectiveness of programs and policies implemented at the schools. They expect to be able to use the scores to plan next year's professional development and to fine-tune their curriculum.

Teachers are admittedly uninformed about the new accountability system. Most say that they do not know very much about the new test, and indeed for the vast majority of teachers CATS and the Kentucky Core Content Tests were either synonymous terms or they had not heard of the Kentucky Core Content Tests at all. Teachers typically receive their information through their school mailbox or from informal conversations with other teachers at the school. The training that has been offered has either not been specific to the new accountability system or has focused on computing the school index rather than the more global implications of implementing the new accountability system.

Teachers have very little confidence that the new accountability system will translate into

better prepared students. They applaud the addition of student-level accountability, but question using transcripts for that purpose. The students they most want held accountable are those for whom transcripts are not important. Teachers want a fairer and more effective system of accountability, but they are not optimistic that CATS represents those improvements.

Teachers are worried that much the same issues that plagued KIRIS will fail to be sufficiently or fairly addressed by CATS. First, the implications of cohort testing continue to be a major point with teachers. They persist in their beliefs that differences between one year's students and the next are sufficient to radically alter a school's index. Second, the accountability system does not sufficiently deal with demographics. Some schools simply have different student populations than others, yet all are measured along the same scale. Worse, the schools least able to improve in the state are the schools that are expected to post the most impressive gains. Third, special education students are included in the assessment, but efforts to ensure the reliability and validity of those students' scores remains questionable. Coupled with the fact that not all schools have an equal proportion of special education students, this issue takes on a great deal of significance at schools with a large special education program.

Other issues exist as well, including a proportion of teachers philosophically opposed to any accountability system. When all of these issues are coupled with the normal debates and concerns surrounding accountability and testing in general, CATS is under a considerable burden to prove its worth. The next few months are of extreme importance for the survival of this newest Kentucky experiment.

The following suggestions may help ensure that CATS is judged on its merit as an accountability system rather than on rumor and misinterpretation:

1. Provide schools with a comprehensive guide for interpreting score reports well in advance of September score report release date. Include appropriate cautions about the divergence from KIRIS. Especially emphasize that the new score reports cannot be used to gauge growth.
2. Describe the new method for calculating a school's index score and "safety zone" in a non-technical school mailing. The need for concise and accurate descriptions of how each school's accountability index will be determined and what that designation means cannot be overstated. Also, since the "safety zones" will not be the same for all schools, this mailing could help ameliorate the confusion schools are likely to have about its determination and function.
3. Provide new teachers with an "Introductory Guide to Kentucky Accountability and Testing" that explains the basics of CATS and the Kentucky Core Content Tests. Their initial year in the state can be overwhelming; a simple guide could greatly improve their knowledge of how their school and students are evaluated in the Kentucky accountability system.
4. Provide incentives for elementary and middle school teachers to bolster their content education.
5. Create opportunities to communicate and cooperate with Kentucky's colleges and universities. Teacher quality is one of the most often cited issues related to improving education. Use this renewed cooperation with higher education to foster improvements in the opportunities for teachers to continue their education within their content area.
6. Evaluate the reliability and validity of testing special education students with and without accommodations. Use those evaluations to gauge the effects of a large special education

demographic on schools and create policy accordingly. Also research the circumstances surrounding certain schools' dramatic increases in special education population.

7. Resolve testing violations quickly and evaluate the effect that penalties based on those violations have on the schools that receive them. Strive for a balance between demanding fair and equal test administration and school punishment.
8. Further research on transience and demographics is clearly indicated. These factors are known to affect test scores and may also affect a school's ability to improve. If schools with high transience or high poverty have less opportunity to improve than other schools, the fairness of the CATS system is compromised.
9. Establish a means of supplementing the DACs during times when communication is particularly necessary.
10. Release information about the new accountability system in a timely manner. Provide the *Kentucky Teacher* and other teacher-oriented publications with non-technical summaries of pertinent research.
11. Research other methods of establishing student-level accountability, paying particular attention to those designed to increase the efforts of lower-achieving students.

This research effort will continue in the coming academic year, 1999-2000. The primary focus of the next phase will be school interpretation and use of the Kentucky Core Content Test score reports as well as a continuing evaluation of the effects CATS is having on public schooling. This next report will also examine aspects of changing perceptions as the new accountability system reaches full implementation.

Bibliography

Abelmann, C. H. & Kenyon, S. B. (1996, April). *Distractions from teaching and learning: Lessons from Kentucky's use of rewards*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Archer, J. (1999, March 31). Out-of-field teaching is hard to curb. *Education Week, 18*(29), pp. 1, 8-10.

Bloom, B. S., Engelhart, M. B., Furst, E. J., & Hill, W. M., & Krathwohl, D. R. (Eds.) (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook I. Cognitive domain*. Ann Arbor, MI: David McKay.

Fishback, F. S. (1998/9, December/January). State board wants teacher feedback on proposed model. *Kentucky Teacher*, pp. 3, 13.

Fishback, F. S. (1999, March). Interim accountability model is on the table for discussion. *Kentucky Teacher*, pp. 1, 3.

Galluzzo, G. R. (1999, May 5). Will the best and the brightest teach? *Education Week, 18*(34), pp. 56, 38.

Hambleton, R. K., Jaeger, R. M., Koretz, D. M., Linn, R. L., Millman, J., & Phillips, S. E. (1995). *Review of the measurement quality of the Kentucky Instructional Results Information System, 1991-1994*. Frankfort, KY: Office of Educational Accountability, Kentucky General Assembly.

Harp, L. (1999, April 13). Board close to approving test system. *The Courier-Journal*, pp. B1, B3.

Harris, C. D., Hoffman, R. G., Koger, L. E., & Thacker, A. A. (1998). *The relationship between school gains in 7th and 8th grade KIRIS scores and instructional practices in mathematics, science, and social studies for 1997* (HumRRO Report FR-WATSD-98-05). Radcliff, KY: Human Resources Research Organization.

Henke, R. R., Geis, S., & Giambattista, J. (1996). *Out of the lecture hall and into the classroom: 1992-93 college graduates and elementary/secondary school teaching* (NCES Report 96-899). Washington, DC: US Department of Education, National Center for Education Statistics.

Hoffman, R. G., Harris, C. D., Koger, L. E., & Thacker, A. A. (1997). *The relationship between school gains in 8th grade KIRIS scores and instructional practices in mathematics, science, and social studies* (LRS97-3). Frankfort, KY: Bureau of Learning Results Services, Kentucky Department of Education.

Hoffman, R. G., Koger, L. E., & Awbrey, A. (1997). *Changes in spelling, capitalization,*

punctuation, and subject/verb agreement skills under the Kentucky Education Reform Act (LRS97-1). Frankfort, KY: Bureau of Learning Results Services, Kentucky Department of Education.

Holland, H. (1998). *Making change. Three educators join the battle for better schools*. Portsmouth, NH: Heinemann.

Ingersoll, R. M. (1999). The problem of underqualified teachers in American secondary schools. *Educational Researcher, 28*(2), March, pp. 26-37.

Kelley, C. & Protsik, J. (1996). Risk and reward: Perspectives on the implementation of Kentucky's school-based performance award program. Paper submitted to *Educational Policy*, May 13, 1996.

Kentucky Department of Education. (1995). *Transformations: Kentucky's curriculum framework*. Frankfort, KY: Author.

Kentucky Department of Education. (1996). *Core content for assessment*. Version 1.0. Frankfort, KY: Author.

Kentucky Department of Education. (1997). *Kentucky consolidated planning process*. Frankfort, KY: Author.

Kentucky Department of Education. (1998a). Request for proposal. Frankfort, KY: Author.

Kentucky Department of Education. (1998b). *Program of studies implementation manuals*. Frankfort, KY: Author.

Kentucky Education Reform Act (Kentucky Revised Statutes 158.645).

Koretz, D. M., Barron, S., Mitchell, K. J., & Stecher, B. M. (1996). *Perceived effects of the Kentucky Instructional Results Information System (KIRIS)*. Santa Monica, CA: RAND.

Lawson, G. (1997, Sept. 11). Adjusted test scores mean 21schools to get cash rewards. *The Courier-Journal*, pp. 1, 8.

McDiarmid, G. W., David, J. L., Kannapel, P. J., Corcoran, T. B., & Coe, P. (1997). *Professional development under KERA: Meeting the challenge*. Lexington, KY: The Partnership for Kentucky Schools and The Prichard Committee for Academic Excellence.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement (Third Edition)* (pp. 13-104). New York: American Council of Education and Macmillan Publishing Company.

Meyer, J. W. & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83*(2), 340-363.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Smith, T. M. (1995). *The educational progress of women. Findings from The Condition of Education 1995* (NCES Report 95-768). Washington, DC: US Department of Education, National Center for Education Statistics.

Stake, R. L. (1995). *The art of case study research*. Thousand Oaks, CA. Sage Publications, Inc.

Thacker, A. A. & Hoffman, R. G. (1998). *School visits and teacher survey: Draft four-year plan*. Radcliff, KY: Human Resources Research Organization.

Thacker, A. A., Hoffman, R. G., and Koger, L. E. (1998). *Course contents and KIRIS: Approaches to instruction and curriculum design in middle school science and social studies* (HumRRO Report FR-WATSD-98-07). Radcliff, KY: Human Resources Research Organization.

Thacker, A. A., Koger, L. E. & Koger, M. E. (1998). *Professional development under the Kentucky Education Reform Act: A study of practices in 30 middle schools* (HumRRO Report FR-WATSD-98-38). Radcliff, KY: Human Resources Research Organization.

The Evaluation Center, Western Michigan University. (1995). *An independent evaluation of the Kentucky Instructional Results Information System (KIRIS)*. Frankfort, KY: The Kentucky Institute for Education Research.

Appendix A
Introduction Letter from the Kentucky Department of Education

January 4, 1999

«Title» «FirstName» «LastName»
«JobTitle», «School»
«Address1»
«Address2»

Dear «Title» «LastName»:

As you may know, the Kentucky Department of Education is sponsoring research on the validity of the CATS assessment and accountability system. The Human Resources Research Organization (HumRRO) is working with KDE to conduct validity studies. These studies include visits to 10 middle schools, 10 elementary schools, and their respective district offices this spring to explore the changes in teaching and learning that are accompanying changes in testing systems. Your school has been chosen to take part in this research.

HumRRO is a non-profit research organization that has been conducting research in instructional practices and testing technology for nearly 50 years. Researchers working on this project have been involved in educational reform issues at the national level and have been conducting similar research in other states. HumRRO is in its third year of conducting studies on Kentucky's assessment and accountability system. HumRRO staff members assigned to the project have the professional experience and capability to efficiently collect critical data and then synthesize those data into meaningful results and interpretations. All staff members participating in this study are Kentucky residents.

This spring, HumRRO plans to continue their investigations in science and social studies. HumRRO researchers will spend one or two days in your school interviewing selected personnel and collecting descriptive information about the school. Researchers will conduct individual interviews with the principal, science teachers, and social studies teachers in the middle schools. They will also observe a portion of the interviewed teachers. We estimate that these interviews will take about 30 minutes each. These may be conducted during planning sessions, before school, or after school. Observations will be conducted at the convenience of the school and will typically last only one class period. Interviewed teachers will be asked to supply copies of three classroom assessments that they have used or will use during the school year.

The data collected during the course of this research will be used to describe how CATS is impacting instruction in representative Kentucky schools. Reports will be submitted to KDE and reviewed by CATS advisory panels as part of the on-going assessment of the validity and credibility of the CATS program. Neither staff members nor schools will be named and every attempt will be made to maintain their anonymity. Unfortunately, because many schools have characteristics that make them unique among Kentucky public schools, complete anonymity cannot be fully guaranteed.

In a few days Art Thacker of HumRRO will be contacting you by telephone to set up a visit schedule and give you more details regarding the study. He is a former Kentucky teacher and will lead HumRRO's effort. If you have questions, feel free to call or e-mail Art in Radcliff, KY at (502) 351-6088 or athacker@humrro.org. If you prefer, you may call Sue Rigney at KDE (502) 564-9855 with your questions.

We appreciate your participation in this research. Our goal is to complete site visits before the CATS testing period this spring. To meet that deadline, we need to complete arrangements and conduct site visits in February, March, and early April. The issues researched as part of this project have the potential to contribute to the body of information on teaching and learning in Kentucky. We look forward to talking to you more about this project.

Sincerely,

Wilmer S. Cody

cc: District Superintendent
District Assessment Coordinators

Appendix B
Letters to Elementary and Middle School Principals

Dear (Principal Name)

Thank you for agreeing to take part in the Kentucky Department of Education's four-year program evaluation of CATS, the state's revised assessment program. The cooperation of educators like you is vital to learning all we can about the educational progress being made in our state.

As I mentioned in our recent phone conversation, a two-member research team will visit your school on (date). We will very likely only be visiting your school for a single day, however if scheduling data collection requires extra time, we might be able to return for a short time the following day. During the visit, researchers will:

- Interview fourth-, fifth-, and sixth-grade teachers. These interviews will last about 30 minutes, and may be done at your teachers' convenience—before or after school or during a preparation period. We would prefer to interview teachers individually, but we understand how hectic schedules are and we will accommodate small groups of teachers if needed. We also understand that teachers are reluctant to give up their sparse planning time and we are committed to limiting individual interviews to 30 minutes. Group interviews may take a little longer.

- Collect three samples of assessment materials from interviewed teachers—lowest unit of assessment (quiz, daily log assignment, etc.), middle unit of assessment (unit test, chapter test, etc), and the largest unit of assessment they use (semester test, grading period test, etc.).

- Observe classroom teaching. In addition to the interviews and the collected material, researchers would like to sit in on some classes. We understand that elementary school may not be divided into clearly delineated classes, but we would like to watch some of the interviewed teachers in action. We will make every effort not to be disruptive and will schedule our observations at the convenience of your teachers. If it is convenient, we would like to see some lessons in science and social studies as part of these observations.

- Interview the principal. This interview could take as long as one hour, but will likely only last about 30 minutes.

We have included several items with this letter to assist you in preparing for our visit.

- Scheduling worksheet. Because our time in your school is limited, we would appreciate your having this worksheet completed before our arrival.

- Letters of introduction for your teachers. Please distribute them at least one week in advance of our visit to give teachers time to collect the requested materials.

- General information sheet, which requests information such as specific directions to your school and recommended motels in your area. We ask that you either mail or fax

this sheet to us as soon as possible to help us in planning our visit.

Please don't hesitate to contact me with any concerns or questions. My telephone number is (502) 351-6088 (call collect) and my e-mail is athacker@humrro.org. Again, thank you for agreeing to take part in this important research. We look forward to meeting you and your staff.

Sincerely,

Art Thacker

Name
Principal
Address

Dear _____,

Thank you for agreeing to take part in the Kentucky Department of Education's four-year program evaluation of CATS, the state's revised assessment program. The cooperation of educators like you is vital to learning all we can about the educational progress being made in our state.

As I mentioned to you during our recent phone conversation, a two-member research team will begin its visit to your school on ____ and ____. One researcher will be on site for two full days. The other researcher will spend the first day at your school, but will visit the district office for about half of the second day. During that time, the researchers will do the following things:

- Interview all science and social studies teachers. These interviews will last about 30 minutes, and may be done at your teachers' convenience—before or after school or during a preparation period. While we prefer to interview teachers individually, we understand that group interviews of teachers of the same subject and same grade may be unavoidable. We realize that teachers are reluctant to give up an entire prep period and we are committed to limiting individual interviews to 30 minutes; however, group interviews will take longer.

- Collect three samples of assessment materials from interviewed teachers—lowest unit of assessment (quiz, daily log assignment, etc.), middle unit of assessment (unit test, chapter test, etc.), and largest unit of assessment (semester test, grading period test, etc.).

- Observe science and social studies classes. Because science is assessed in seventh grade and social studies in the eighth grade, we would like to ensure that we are able to observe at least a few of those particular classes. Observations in non-assessed subjects (for example, seventh grade social studies) would also be useful, and we would like to observe as many of those classes as possible. Should scheduling conflicts develop, we prefer that interviews receive priority rather than observations.

- Interview the principal. This is estimated to take about one hour.

We have included several items with this letter to assist you in preparing for our visit.

- Scheduling worksheet. Because our time in your school is limited, we would appreciate your having this worksheet completed before our arrival.

- Letters of introduction for science and social studies teachers. Please distribute them at least one week in advance of our visit to give teachers time to collect requested materials.

- General information sheet, which requests information such as specific directions to your school and recommended motels in your area. We ask that you either mail or fax this sheet to us as soon as possible to help us in planning our visit.

Again, thank you for agreeing to take part in this important research. We look forward to meeting you and your staff.

Sincerely

Art Thacker

Appendix C
Elementary and Middle School
Worksheet Scheduling Forms

Scheduling Worksheet

We have included a worksheet (on the following page) to help you schedule interviews and observations which will take place during our visit to your school. You will have two researchers available. As we previously mentioned, we would like to interview the following people:

- The principal
- All fourth-, fifth-, and (where available) sixth-grade teachers

Please write in the name of the person to be interviewed, along with the time and place of the interview in the appropriate cell. We have found that interviews go more smoothly if they are conducted in a quiet place (for example, an empty classroom or conference room, corner of the media center). Also, please remember that although two interviews can take place at the same time, they should not be scheduled for the same location in order to respect the privacy of the person being interviewed.

We would like to observe the following classes:

- Priority to fourth- and fifth-grade classes. Whenever possible we'd like to observe fourth-graders doing science and fifth-graders doing social studies.
- Other observations as available and as time permits.

Please keep the following in mind when completing the schedule:

- Both researchers will be available during their one-day visit to your school
- Two interviews can be scheduled for the same time, but not for the same place
- An individual interview will last about 30 minutes, group interviews may last a little longer
- If scheduling conflicts develop, please keep in mind:

1. Principal and fourth- and fifth-grade teacher interviews receive top priority
2. Fourth-grade science and fifth-grade social studies observations receive next priority
3. All other observations and interviews receive final priority

- Our time in your school is limited. Therefore, it is very important that we use that time as efficiently as possible. We would appreciate it if you would complete this schedule before our arrival.

(Note; the other category refers to any teacher who you feel might be appropriate for interview/observation in order for us to more accurately depict your school, for instance if your school has a science or social studies specialist that does not fit neatly into a specific grade level, etc.)

Elementary School Scheduling Worksheet. When completing this worksheet. Please remember that **two** researchers will be visiting your school. The team can interview two teachers in 30 minutes or observe 2 one-hour classes at the same time. Since they have only one day to complete the visit, they will rely heavily on your help in scheduling these events.

Interviews

	1 Principal	1-3 4th Grade Teachers	1-3 5th Grade Teachers	1-3 6th Grade Teachers	Other
Name Time Interview Place					
Name Time Interview Place					
Name Time Interview Place					

Observations

	4th Grade	5th Grade	6th Grade
Name Time Classroom			
Name Time Classroom			
Name Time Classroom			

Middle School Scheduling Worksheet. When completing this worksheet, please remember that **two** researchers will be visiting your school. The team can interview two teachers in 30 minutes or observe two one-hour classes at the same time. Since they have only about one and one half days to complete the visit, they will rely heavily on your help in scheduling these events. Please try and schedule all interviews during the first day.

Interviews

	1 principal	**1-3 6th science teachers**	**1-3 7th science teachers**	**1-3 8th science teachers**	**1-3 6th social studies teachers**	**1-3 7th social studies teachers**	**1-3 8th social studies teachers**
Name Time Interview place							
Name Time Interview place							
Name Time Interview place							

Observations (If possible, please give priority to 7th-grade science and 8th-grade social studies.)

	6th Science	**7th Science**	**8th Science**	**6th Social Studies**	**7th Social Studies**	**8th Social Studies**
Name Time Classroom						
Name Time Classroom						
Name Time Classroom						

Appendix D
Travel Arrangement Form

Please fax or mail the following information:

TO: Art Thacker

FAX: (502) 351-3620

MAIL: 295 W. Lincoln Trail Blvd.
Radcliff, KY 40160

FROM:

1. In addition to the principal, who else may we contact, particularly if we need to make last-minute adjustments for bad weather? Are there alternate phone numbers available?

Name	Alternate Phone Number

2. Please give us directions to your school or sketch a map:

3. Do we need to know anything about parking restrictions at your school?

4. Can you recommend a motel near your school? (Name, location, phone number)

5. What time does your school day begin (what time do we need to arrive)?

Appendix E
Elementary and Middle School
Teacher Information Letters

Information Sheet for Elementary Teachers
Explanation of HumRRO Research

Background

For the past two years, middle schools in Kentucky have taken part in research studies conducted by the Human Resources Research Organization (HumRRO). The first study, in 1997, investigated how teaching practices were related to KIRIS scores. Researchers found evidence that use of reform practices, such as cooperative learning, extended problem solving, discussion, and student writing, were more likely to increase KIRIS scores than were memorization-based practices. The 1998 study examined how schools use professional development and other teacher preparation activities to adapt to the demands of Kentucky's educational goals.

This year, we plan to expand our research efforts by visiting 10 elementary schools. We plan to examine the impact of changing the statewide assessment system from KIRIS to CATS. Your school has agreed to take part in this study, which will run through 2002. The first of your four visits will take place on March 18. We are asking for your help in a couple of ways:

- We plan to interview all fourth-, fifth-, and sixth-grade teachers at your school, concentrating on science and social studies instruction. Each interview will last about 30 minutes, and can be done at your convenience—before or after school or during a planning period. We would prefer to conduct individual interviews, but we realize how hectic schedules are and we will accommodate small groups of teachers if necessary. We also understand that planning time is in short supply and we are committed to limiting individual interviews to 30 minutes. Group interviews may take a little longer.

    - We request that you bring three samples of assessments that we may keep to your interview. These should represent three different levels of assessment: a basic unit of assessment, such as a quiz or log book assignment; a mid-level unit of assessment, such as a chapter test, unit test, etc.; and an upper-level unit of assessment, such as a semester test, end-of-grading-period test, etc. Actual samples of student work are not required.

    - We also would like to observe classroom teaching. We understand that elementary school may not be divided into clearly delineated classes, but we would like to watch some of you in action. We will make every effort not to be disruptive and will schedule our observations at your convenience. If possible, we would like to see some lessons in fourth-grade science and fifth-grade social studies.

Confidentiality

We will not identify participating schools or personnel in any report, presentation, or discussion of this research. No information collected by interview, observation, or conversation will be divulged to any administrator, teacher, staff, or student within your school, or to any Kentucky Department of Education staff member. Written reports will provide information in summary form only. However, because your school may have characteristics that make it unique among Kentucky schools, anonymity cannot be guaranteed. If you or other members of the school staff read the report, you may be able to determine that certain parts refer to your school. It is highly unlikely that anyone less familiar with your school would recognize it in the report. Please remember that this research is being conducted to evaluate CATS, not schools.

HumRRO Contacts

Please contact Art Thacker at (502) 351-6088 (email address: athacker@humrro.org) if you have any questions or concerns.

Information Sheet for Science and Social Studies Teachers
Explanation of HumRRO Research

Background

For the past two years, middle schools in Kentucky have taken part in research studies conducted by the Human Resources Research Organization (HumRRO). The first study, in 1997, investigated how teaching practices were related to KIRIS scores. Researchers found evidence that use of reform practices, such as cooperative learning, extended problem solving, discussion, and student writing, were more likely to increase KIRIS scores than were memorization-based practices. The 1998 study examined how schools use professional development and other teacher preparation activities to adapt to the demands of Kentucky's educational goals.

This year we plan to visit 10 middle schools to examine the impact of changing the statewide assessment system from KIRIS to CATS. Your school has agreed to take part in this study. The first of four visits to your school to take place between now and 2002 is scheduled for March 22 and 23. We are asking for your help in a couple of ways:

- We plan to interview all science teachers and social studies teachers at your school. Each interview will last about 30 minutes and will be held at a time convenient for you (before/after school or during your planning period). We prefer to interview teachers individually, but we realize that may not be possible in some situations. If a group interview is necessary, we ask that it consist only of teachers who teach the same subject in the same grade. Group interviews will take longer than 30 minutes.

- We request that you bring three samples of assessments to your interview that we may keep. These should represent three different levels of assessment: a basic unit of assessment, such as a quiz or log book assignment; a mid-level unit of assessment, such as a chapter test, unit test, etc.; and an upper-level unit of assessment, such as a semester test, end-of-grading-period test, etc. Actual samples of student work are not required.

- We also plan to observe science and social studies classes. Our first priority will be seventh-grade science and eighth-grade social studies, since those subjects are assessed in those grades. However, as time permits, we would like to observe classes in other grades as well.

Confidentiality

We will not identify participating schools or personnel in any report, presentation, or discussion of this research. No information collected by interview, observation, or conversation will be divulged to any administrator, teacher, staff, or student within your school, or to any Kentucky Department of Education staff member. Written reports will provide information in summary form only. However, because your school may have

characteristics that make it unique among Kentucky schools, anonymity cannot be guaranteed. If you or other members of the school staff read the report, you may be able to determine that certain parts refer to your school. It is highly unlikely that anyone less familiar with your school would recognize it in the report. Please remember that this research is being conducted to evaluate CATS, not schools.

HumRRO Contacts

Please call Art Thacker at (502) 351-6088 (email address: athacker@humrro.org) if you have any questions.

Appendix F
Teacher Interview Protocol

Hello, my name is _____. I am a researcher with Human Resources Research Organization (HumRRO). HumRRO is a private, non-profit research organization under contract to the Kentucky Department of Education (KDE).

First, have you had a chance to review the information sheet about HumRRO and what we are doing? ***(show information sheet to the teacher)***

> ***(If the teacher has seen the information sheet ask***
> Do you have any questions? May I collect your assessment samples now?
>
> ***(Read this section only if the teacher is unfamiliar with the information sheet or asks for more information.)***

HumRRO's Task

(Skip if teacher has read info sheet)

Our task is to collect and analyze data on validity issues associated with the Commonwealth Accountability Testing System (CATS) and Kentucky Core Content Test (KCCT). We are independent from the test developers and report directly to the Kentucky Department of Education (KDE) with our findings. We have worked with KDE investigating validity issues with KIRIS for the past two years. This is the first of four years we will be investigating the validity of the new tests.

Thanks
(Begin again here)

Thank you for taking the time to complete this teacher interview.

Objectives

We want to hear

- what you think about the new accountability and testing system and
- how you believe it will affect you and your students.

We also want to get a better understanding of

- your instructional approaches and
- activities you use in your classroom.

Perceptions

I will be asking for your perceptions about CATS. I understand that you have not seen the test or the reports yet and you may or may not have been told very much about them.

Confidentiality	• We will not identify participating schools or personnel in any report, presentation, or discussion of this research. • However, because your school may have characteristics that make it unique among Kentucky schools, anonymity cannot be guaranteed. If you or other members of the school staff read the report, you may be able to determine that certain parts appear to refer to your school. It is highly unlikely that anyone less familiar with your school would recognize it in the report.
Target of Evaluation	Please remember that this research is being conducted to evaluate CATS, not teachers, schools, or districts.
Length of Interview	The interview was designed to take less than 30 minutes.

Do you have any questions before we begin?
(If not, turn to next page to begin interview.)

(Anticipated questions that we will need set answers for)
When will this report be finished?
Who will see this report?
Will we get copies of it?
Is it true that you'll be coming back in following years?
How was our school selected?
Are you visiting other schools in our district?
Why were just science and social studies teachers selected?
What will you do with these interviews?
Have you seen the CATS test yet?

The first section is designed for us to collect some background information on you.

- ***Note: Researcher will read back what you have recorded for verification.***

1. What subject and/or grade(s) do you teach this year?	
Subject	Grade

2. How long have you taught this (these) grade(s) and/or subject(s)?	
Grade(s)	Subject(s)

Have there been any reorganizations within the school in the last four years? (Such as some grades moving to another school, changing to team teaching, block scheduling, or redistricting. This does not include getting a new administration at the school.)

3. Counting this year, how long have you been teaching at this school?	

4. Counting this year, how many years have you been teaching in Kentucky? Elsewhere?	In Kentucky	Elsewhere

This section of the interview deals with your perception of KIRIS and CATS. We are interested in finding out what you think about the systems and what effect you believe CATS will have on how you teach and what students learn.

1. How much have you heard about CATS?
(follow up with; Where did you learn about CATS?)

2. Do you think CATS will be better or worse than KIRIS? Why?

3. Do you think the new system will have an effect on how and what students learn?

(if "yes") What are the major differences in learning measured by CATS?

4. How confident are you that CATS will help the state administer rewards or assistance more fairly?

(ask follow up question such as "Why do you feel that way?")

5. Do you think that students will be better prepared for

(elementary teacher) middle school
(middle school teacher) high school

after CATS compared to before CATS?

6. Do you think your school's chances of receiving awards increase or decrease under CATS?

Tell me why you think your school's chances are better/worse.

7. If scores go on student transcripts, do you think students here will take CATS tests less or more seriously than the KIRIS tests?

8. How useful were the 1997-98 KIRIS scores compared to score reports from previous years?

The next section of the survey deals with your **teaching methods or instructional**

practices.

1. Do you think you have changed the way you teach because of the change in testing from KIRIS to CATS?

(If the teacher indicates changing ask the following)
Please indicate about how much you have changed the way you teach because of the change in testing.

Please tell me about some of the changes you have made. *(Make sure that if the teacher lists several changes to ask which is most important.)*

2. Do you expect the addition of the multiple-choice questions to the school accountability formula to affect teaching here? In what way?

What will that mean in terms of student learning?

The last section asks about teacher professional development.

1. Has CATS had any influence on the professional development you have had this year? Better or worse than before?

(Why is it better or worse?)

Appendix G
Principal Interview Protocol

Principal Interview 2 (abbreviated version)

1. Tell me about this year's preparation for CATS. What specific things were done at the school to get teachers and students ready?

2. Is that any different from last year's preparation for KIRIS? How?

3. How were the 97-98 KIRIS score reports used this year?

4. How useful were they? What would have made them more useful?

5. Will the changes in the way rewards are structured affect teacher motivation? (School

rewards rather than personal bonuses) Comment/Explain?

6. Have you seen any changes in teaching practices associated with switching from KIRIS to CATS? Describe them.

7. The multiple choice questions on CATS are going to count this year and a norm-referenced multiple choice test is being added as a component to school accountability scores. Will that have any effect on teaching and learning here? Explain/Describe?

8. Will the addition of test scores on student transcripts cause the students here to take the test more seriously? Why or why not?

9. What one school-wide project, program, policy, etc., implemented here, has had the most influence on students and teachers? Describe it. (When implemented, who implemented, term length, unanticipated benefits/consequences)

10. *Elementary Teachers Only:* Describe the influence of the middle school on the curriculum/teaching practices here at this school, if any. Describe the influence of the district, if any.
    Middle School Teachers Only: Describe the influence of the elementary school on the curriculum/teaching practices here at this school, if any. The high school? Describe the influence of the district, if any.

Appendix H
District Interview Protocol

District interview

1. What are some of the more important things the district does to help schools prepare for CATS/KCCT?

2. Has the district office's role changed as a result of the switch from KIRIS to the KCCT? How?

3. Does the district sponsor any professional development aimed specifically at getting schools/teachers ready for the KCCT? (*Skip if PD is already mentioned as part of 1 or 2.*)

4. What issues related to CATS seem to garner the most concern among schools within the district?

5. Has the addition of the NRT and multiple choice scores to the school accountability index caused any noticeable changes in the schools here in the district? Describe the changes.

6. How does the district office promote cooperation/communication between the schools within the district? What might make those efforts more successful?

7. How did the district use the 1997-98 KIRIS reports? What kind of district-level decisions will be made using the CATS reports? From what you know about the proposed structure of the reports, which will be more useful; KIRIS or CATS? Why?

8. Rewards are to be administered as school funds rather than as teacher bonuses. What effect will that have on teacher/school motivation to improve?